DOCUMENT RESUME

ED 319 116 EA 021 798

AUTHOR Cuadra, Ernesto; Crouch, Luis
TITLE Indicators of Student Flow Rates in Honduras: An Assessment of an Alternative Methodology, with Two Methodologies for Estimating Student Flow Rates. BRIDGES Research Report No. 6.
INSTITUTION Harvard Univ., Cambridge, MA. Inst. for International Development.
SPONS AGENCY Agency for International Development (IDCA), Washington, DC.
PUB DATE Dec 89
NOTE 45p.; Project directed under Cooperative Agreement No. DDP-5824-A-5076.
AVAILABLE FROM Publications, Project BRIDGES, 442 Gutman Library, Harvard University, Cambridge, MA 02138 ($5.00).
PUB TYPE Reports - Descriptive (141) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Access to Education; *Data Collection; *Developing Nations; Educational Planning; Educational Quality; *Efficiency; Elementary Secondary Education; *Enrollment Projections; Foreign Countries; *Grade Repetition
IDENTIFIERS *Honduras

ABSTRACT

Student promotion, repetition, and dropout rates constitute the basic data needed to forecast future enrollment and new resources. Information on student flow is significantly related to policy formulation aimed at improving internal efficiency, because dropping out and grade repetition increase per pupil cost, block access to eligible school-age children, and prevent students from gaining an adequate education. Although studies conducted in many Third World countries have shown that student enrollment information collected by education ministries is fairly reliable, most countries underestimate grade repetition rates, because reliable data is absent. One method for calculating the number of repeaters that is not based on school data is the age/grade method. The basic assumption is that enrollment data by age and grade sent to education ministries is more accurate than the repetition data. Cuadra's paper describes a 1986 study conducted in Honduras that tested this model by comparing school records of 327 students with data obtained from interviewing the students' parents. The study determined the accuracy of the age/grade method for reporting basic enrollment information and found that school statistics consistently underreported the number of repeaters. By underreporting repeaters, schools also overestimated the number of dropouts. Other implications are discussed. Two other papers discuss alternative methods for estimating student flow rates in the absence of sufficient data. (52 references) (MLH)

ED319116

Methodology

by Ernesto Cuadra

with

Two Methodologies for Estimating Student Flow Rates

by Luis Crouch

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

EA 021 796

B • R • I • D • G • E • S

Basic Research and Implementation in DevelopinG Education Systems

A project of the Harvard Institute for International Development,
the Harvard Graduate School of Education,
and the Office of Education, Bureau for Science and Technology,
United States Agency for International Development

The Basic Research and Implementation in DevelopinG Education Systems Project (BRIDGES) is directed by the Harvard Institute for International Development and the Harvard Graduate School of Education, under Cooperative Agreement No. DDP-5824-A-5076 with the Office of Education, Bureau for Science and Technology, United States Agency for International Development. Also participating in the Project are the Institute for International Research, Michigan State University, the Research Triangle Institute, and Texas Southern University.

The BRIDGES Group includes educators, researchers, planners and policymakers committed to improving opportunity and quality in Thirld World schools. The goal of their collaborative effort is to identify policy options that will increase children's access to schooling, reduce the frequency of early school leaving and repetition, improve the amount and quality of what is learned, and optimize the use of fiscal and educational resources.

The *BRIDGES Research Report Series* is edited by the Harvard Institute for International Development. The *Series* is a collection of reviews of the state-of-the-art in research, and original research reports on basic education in developing countries. Each review summarizes research about a particular policy issue and suggests policy options. Original reports on BRIDGES-sponsored research present new information about the impact and costs of specific alternatives that the reviews have identified as most promising.

The views expressed in this document are those of the author and do not necessarily reflect those of the United States Agency for International Development.

Indicators of Student Flow Rates in Honduras: An Assessment of an Alternative Methodology

Ernesto Cuadra

with

Two Methodologies for Estimating Student Flow Rates

Luis Crouch

Basic Research and Implementation in DevelopinG Education Systems
(B R I D G E S)

A Cooperative Agreement Between Harvard University and the Office of Education, Bureau of Science and Technology, with the assistance of the Office of Development Resources, Bureau for Latin America and The Caribbean. United States Agency for International Development. Project number DPE-5824-A-00-5076-00

ACKNOWLEDGEMENTS

This study was a cooperative effort of BRIDGES, the Ministry of Education in Honduras, and the Bureau for Latin America and the Caribbean of the United States Agency for International Development.

Valuable assistance in the design of the study, the coordination of the filed work, and the coding and preliminary analysis of the data was provided by Orlando Betancourth. Jose Roberto Martinez and Raul Mungula, from the Research Unit of the Ministry of Education, greatly assisted in the data collection, coding, and preliminary analysis of the data. Vilma Daysi Guillen, Sagrario Galdo, Estela Valentina Paz Bonilla, Herlinda Maribel Molina, Enrique Burgos, Darwing E. Guillen, Amadeo Juarez Osorio and Alejandro Arita, from the In-Service Teachers' Training Unit of the Ministry of Education, conducted the pilot test, assisted in redesigning the instruments for data collection, collected the data from schools, and conducted parent interviews.

Josefina Gamero, Director of the In-Service Teachers' Training Unit, Marco Tulio Mejia, from AID-Honduras, Omar Rousel, Liason between the Ministry of Education and AID-Honduras, and Carleton Corrales, Director of the Teachers' College, generously gave of their time to discuss the first draft of the research design and gave valuable advice in the selection of the schools where the data were collected.

Special thanks to Joy Wolf Shepard and Pam Summa for their editorial assistance.

Foreword

Educational planners and policymakers face tough choices. In many systems the quality of schooling is below minimal standards, and not all children have gained access to school. Because resources are limited, increased spending on improvements in quality will require reduced spending on enrollment expansion. The quantity versus quality choice is one that many countries feel obliged to make.

There is evidence, however, that the total resources available to expand quantity or to raise quality can be increased significantly through improvements in the efficiency of the education system— that is, without increases in total spending. An improvement in the efficiency of an education system increases the use of resources that are already available, either to expand quantity or to improve quality, with no additional provision of resources.

A major symptom of inefficiency in schools is student repetition, which occurs when the education system decides that a child has not been adequately prepared for the next grade. If schools were fully efficient, the instruction offered during one grade would be enough to prepare all students for the next grade. Some educators cite dropouts, or early school leavers, as a more important symptom of inefficiency. If a student leaves the system before completion of a cycle (e.g., before completing all grades of primary school) the effort put into the education of that child is considered wasted. This view is supported by evidence that graduates of a 6-year primary cycle earn much more than students who have completed only 5 years, and that there are small differences in earnings between school leavers with 3 or 4 or 5 years of primary school. As a consequence, dropout prevention, in some countries, is given more attention than reduction of repetition. Large sums are spent on programs for children who have left school early.

In most cases, as Ernesto Cuadra argues, it is a mistake to focus primarily on the problem of dropouts. It is important to reduce the dropout rate, but repetition is actually a more important problem to tackle. In fact, repetition rates in any given grade are generally higher than dropout rates. Further, research shows that repetition is the main cause of dropping out, and that reduction of failure and consequent repetition leads to reduced dropout rates, and higher rates of completion of the school cycle. Because repeaters take up seats that could be occupied by new students, reduction of repetition also increases the space available in the system without additional capital expenditures. This is particularly important in the lower grades of the primary cycle, where very high repetition clogs the system, preventing new students from entering.

Planners may fail to recognize the limitations on access caused by high repetition. Conventional statistical methods underestimate repeater rates and exaggerate dropout rates. As a consequence the conventional methods also overestimate the number of students who enter the first grade for the first time each year. As Cuadra demonstrates in this paper, underestimating repetition in the early grades can create the illusion that a country is making headway in providing education to all children, when in fact it is merely requiring some students to spend multiple years in the same grade.

Research shows that repetition is the main cause of dropping out...

Cuadra goes on to demonstrate, through field research in Honduras, the limitations of the conventional method for estimating repetition rates. He notes the conditions under which teachers are likely to supply artificially low estimates of repetition. Because dropout rates and the number of new entrants are calculated from repetition rates, they will be overestimated. He reviews a method that uses information about enrollments by age and grade which can provide highly precise estimates of all three measures.

Not all countries currently collect information on student ages by grade. The two methodologies in the Appendix, by Luis Crouch, offer suggestions for how to estimate repetition rates when age/grade data are

not available. These methods also are superior to methods that rely on teacher reports of repetition.

This set of papers carries the happy news that planners need not regard their choices as being between expansion of quantity, or improvement of quality. Because repetition affects both dropping out and access, and is in turn a consequence of low quality (i.e., student failure), efforts to improve educational quality can also improve the internal efficiency of education systems, freeing up resources that can then be used to provide more space for new entrants. Improved quality can lead to increased quantity, if the focus is on efficiency.

Noel F. McGinn
July 28, 1989

Table of Contents

Executive Summary 2

Section I: Introduction 4
- Purpose of the Paper 4
- Changes in Educational Planning Practices 4
- Role of Information in Improving Internal Efficiency 5

Section II: Overview of the Quality of Educational Information in Third World Countries 7
- Reliability of Student Flow Data 7
- Description of the Grade/Retention Method 8
- Assessment of Student Flow Data in Honduras 8

Section III: An Alternative Approach to Producing Reliable Estimates of Student Flow Data 11
- Description of the Age/Grade Method 11
- Comparison of the Age/Grade Method with the Grade/Retention Method 12
- Application of the Age/Grade Method to Estimate Flow Rates in Honduras 12

Section IV: Empirical Evaluation of the Age/Grade Method 14
- Research Objectives and Data Collection Strategy 14
- Description of the Sample 15
- Data Collection 15
- Data Analysis 16
  - Validity of Data on Age and Grade 16
  - Validity of Flow Data 17
- Conclusion 19

Section V: Factors Affecting Data Accuracy 21

Section VI: Recommendations for Improving Educational Data 23

Appendix 25

Bibliography 28

Two Methodologies for Estimating Student Flow Rates 31
- A New Approach to the Estimation of Enrollment Transition Rates: Estimating Repeater and Dropout Rates with Minimal Data Availability 31
- Monitoring Internal Efficiency of a School System in the Absence of Reliable Repeater and Dropout Data: A Method Based on Simple Proxy Indicators 36

Executive Summary

Educational planners and policymakers need information about how students progress through the school system to assess the efficiency of the system, to forecast future enrollments, to estimate the amount of new resources that the system will need to accommodate those enrollments, and to estimate the cost associated with it. Student promotion, repetition, and dropout rates constitute the basic data needed to forecast future enrollment and new resources. These rates, together with achievement test results, are also the two outcomes of the education system most widely used to evaluate the internal efficiency of the system.

How much students learn and how many students drop out, repeat, or successfully complete a grade (student flow) are all dependent upon policies that determine the amount of direct and indirect resources assigned to education, and the administrative regulations regarding promotion. The design and implementation of successful policies in these areas are dependent on accurate information about students' achievement and how students progress through the school system. Information on student flow is of particular significance in the formulation of policies aimed at improving internal efficiency because dropout and repetition increase per pupil cost, block access to eligible school-age children, and prevent students from gaining the basic education that would provide social and economic opportunities later in life.

To record the number of repeaters, schools must rely on the teachers' recollection of who was enrolled in each grade during the previous year, or on students' own reporting.

Studies conducted in many Third World countries have shown that the information collected by ministries of education from schools regarding student enrollment is fairly reliable. The information on student repetition, however, tends to be less reliable. For example, a research team from BRIDGES visiting six Central American countries found that repetition rates were sometimes underestimated by more than 100 percent. In Nepal a study conducted by New ERA found that schools generally did not keep records on students' progress. BRIDGES found the same situation in the Yemen Arab Republic. In fourteen Latin American countries Schiefelbein (1975) found that, with the exception of Uruguay, all the countries underestimate repetition rates.

The reason that repetition rates are not accurate is because many schools do not require the students to show their school certificate when they enroll at the beginning of the academic year. To record the number of repeaters, schools must rely on the teachers' recollection of who was enrolled in each grade during the previous year, or on students' own reporting. Sometimes students do not even know what it means to be a repeater, especially students in lower grades.

One method for calculating the number of repeaters that is not based on school data is the age/grade method. This method instead uses the school data on enrollment by age and grade for two consecutive years to estimate the total number of repeaters by grade. The basic assumption of this method is that the data on enrollment by age and grade that schools send to the ministry of education is more accurate than the repetition data.

The age/grade method was used to design a simulation model for estimating promotion, repetition, and dropout rates in six Central American countries. In all six countries the model produced repetition rates that were higher than the official estimates published by the ministries of education (MOE). The MOE rates were based on school repetition data collected at the beginning of the academic year.

The model of the age/grade method was tested in a study conducted in Honduras in 1986. This study examined school records on enrollment by grade, age, and promotion patterns for a cohort of 327 students who entered school for the first time in 1983, and compared this information with data obtained from interviews with the students' parents.

The study showed that there was almost 90 percent agreement between school records and parents on both the children's age and the grade attended during each of the four years of the study. This indicates that the basic information used by the simulation model, which is based on the age/grade methodology, is fairly accurate.

The study not only determined the accuracy of the age/grade method, it also compared school records to parents' information on repetition, and discovered that the school's statistics consistently underreported the number of students repeating a grade. For example, in 1984 schools reported that only 63 out of 327 students (or 19 percent) repeated first grade. Parents reported that 144 students (or 44 percent) repeated first grade. In 1985, school records reported 44 percent fewer repeaters than parents; in 1986, 45 percent fewer repeaters.

As a consequence of underreporting repeaters, schools overestimated the number of dropouts. Since schools do not keep records of who drops out of school, dropouts are estimated as the difference between the number of students promoted to the next grade and the number of students who repeat. Based on the school records on the number of repeaters, the number of dropouts at the end of 1986 was 152 students, or 46 percent of the students who entered first grade in 1983. According to the parents' information only 35 students (or 11 percent) had dropped out by 1986.

The implications of this study are that given the importance that having accurate information on students' flow rate has for educational planning, forecasting, and the definition of educational policies, ministries of education should pay more attention to how repetition data are collected by schools and should try to improve the quality of the data that schools send to their statistics offices. Since the improvement in data quality will probably take some time, and since ministries need student flow data for their educational forecasting and planning, ministries will be better off using the age/grade method to estimate promotion, repetition, and dropout. Because computer technology has improved to the point where computers are relatively cheap and easy to use, it is now feasible to use methodologies such as the age/grade method that would be very demanding to compute manually. In countries where information on enrollment by age and grade is not collected on a yearly basis this method cannot be used, and the only wide-ranging solution is to improve the data collection process.

Section I: Introduction

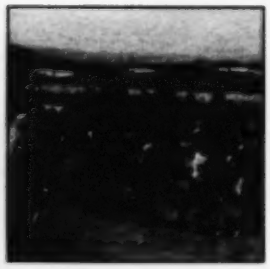
Purpose of the Paper

The present reduction in the amount of new resources available for education, combined with the considerable technological progress in information technologies, has had a significant impact on how educational policymakers and planners value and use information.

Under conditions of growth and stable funding there was little pressure to use information. If a decision maker missed an opportunity to improve the educational system, it was likely that the option to do so would remain open. Under austerity the situation changed dramatically. Today, the allocation of resources and investment demands a more careful evaluation of options as well as of the effects of each decision on different social groups. Information has become a key resource needed to support decisions and to define policy options.

The evolution in computer technology towards more inexpensive, easy to use, and standardized equipment has made the development of computer-based management information systems more accessible to planners and policymakers. These people not only have ready and easier access to more information but also to more sophisticated forms of data analysis and data presentation than ever before.

Unfortunately, the attention given to providing easy access to more information and to more powerful techniques for data analysis has not been matched by a similar attempt to ensure the quality of the data collected and reported by ministries of education (Chapman and Boothroyd, 1988). To improve the quality of data will take time and money; however, until this is done policymakers and planners can use computer simulation models to get a more accurate picture of their education systems.



The attention given to... more powerful techniques for data analysis has not been matched by... the quality of the data collected...

This paper presents the results of a study conducted in Honduras to test one such model. This model was designed to estimate repetition rates by using information on enrollment by age and grade. These rates are then used to project student enrollment and to assess the efficiency of the system.

Changes in Educational Planning Practices

Both total educational enrollments and gross enrollment ratios[1] have expanded continuously, worldwide, since 1950. During that time, the rate of increase for Third World countries has been much higher than for developed countries. As a result, Third World countries' share of total world enrollment grew from 45 percent to 63 percent between 1960 and 1980 (Coombs, 1985, p.74). However, in spite of the large increases in enrollment the majority of these countries have still not reached the level of educational participation found in developed countries (UNESCO, 1983).

As enrollment in educational systems increased, resources allocated to education also grew, especially prior to the mid-1970s. Public expenditure on education increased both as a percent of the national budget and a proportion of the gross national product (GNP). The rate of increase in public expenditures on education between 1960 and 1982 was higher in Third World countries than in developed countries (Schiefelbein, 1983; Lewin, 1987). However, developed countries still commit a larger fraction of both their national budget and their GNP to education than do Third World countries.

In spite of continued growth in the demand for education, the rate of increase in expenditure slowed during the 1970s as a consequence of the world economic recession. According to Schiefelbein (1983), public expenditures on education, as a percent of GNP, rose by 43 percent in developed countries between 1960 and 1970 and by 48 percent in Third World countries. However, between 1970 and 1977,

these figures increased by only 12 percent in developed countries and by 21 percent in Third World countries.

The reduction in the rate of growth in expenditures on education has had a greater impact on Third World countries than on developed countries, both in respect to the public sector's capability to satisfy the increasing demand for education and in its ability to maintain educational quality at its current level (Lewin, 1987 and 1988). Third World countries have experienced a reduction in their already low capabilities to build new classrooms, hire new teachers, improve curriculum, distribute free textbooks, and maintain current facilities. Although developed countries have also faced reductions, their educational systems have a higher level of student participation, which lessens the pressure to expand access.

The combination of continuing growth in the demand for education and the reduction in the rate of growth in new resources available for education has altered the climate in which policy decisions are made. Economic uncertainty, instability, and foreign debt have replaced expectations for economic growth and stability. This change has had a significant impact on the practice of educational policymaking, for "planning within stable or declining budgetary positions is not simply the opposite of planning under conditions of growth" (Lewin, 1988, p. 11). Budgetary constraints not only affect the amount of resources available for education but also bring about organizational changes, goal displacement, redefinition of priorities, greater uncertainty, and the need for rapid decisionmaking. Since small savings in some areas may do more to lower the quality of schooling than bigger cuts in other areas, educational policymakers must reappraise the working practices of their institutions before introducing any cut.

An analysis of the educational goals of more than 125 countries between 1950 and 1965 (Fiala and Gordon, 1987) found that, during the so-called "educational revolution" of the 1950s and 1960s, there was a widespread belief shared by many political leaders, consultants, and intellectuals that countries should invest heavily in education in order to achieve economic growth.[2] Fiala and Gordon call this belief "the ideology of education". Regardless of whether or not this notion is valid, it gave new support to the expansion of educational opportunities. The "ideology of education" not only provided a rationale throughout the world for investing in education, it also promoted specific models of educational organization and management (Inkeles and Sirowa, 1983; Fiala and Gordon, 1987). It was during this period that donor organizations and lending agencies made the introduction of educational planning a condition for assistance (Weiler, 1985; Khoi, 1985).

Lewin (1987 and 1988) characterizes the organizational culture in which educational planning took place between the 1950s and the mid-1970s as "laissez faire." He uses the term to refer to a situation where optimistic expectations of economic growth and stability reduced the need for rapid decision making. Planners operated under the assumption that resources for expansion, innovation, and experimentation in education would be available.

Due to the effects of a global economic recession on the economies of Third World countries and the continuing demand for educational services, the laissez faire climate of the 1950s and 1960s has disappeared. Initially it was replaced by the practice of planning for crisis, in which the dominant force was an uncertainty about the future within a context of shrinking resources. Later, under pressures of mounting foreign debt payments, increasing energy costs, and falling commodity prices, planning for crisis was gradually replaced by planning for budget cuts. Planning that is based upon an assumption that cuts will and must be made has been both extremely conservative, regarding the introduction of innovations, and unrestrained, regarding management practices oriented toward rationalization and cost reductions (Coombs, 1985; Lewin, 1987 and 1988).

Role of Information in Improving Internal Efficiency

The interest in rationalization and cost reduction has focused attention on the internal efficiency of educational systems as the central issue in Third World countries. This concern is not new among managers, planners, and policymakers working in education, nor is it unique to Third World countries (Murnane, 1987; Lockheed and Hanushek, 1988). However, due to the recent concern for austerity and the prevalence of planning for budget cuts, there has been a shift in the emphasis of the argument used by educational planners and consultants to explain why it is important to improve the internal efficiency of the school systems in Third World countries.

Until the early 1970s, efficiency in educational systems was perceived as a means for providing expanded access to education (Brown, 1966; UNESCO, 1967 and 1977; McGinn and Davis, 1969; Levy, 1971;

IBE, 1971 and 1972; Schiefelbein, 1975; Dominguez, 1980). An increase in internal efficiency was considered to be the primary way to expand educational opportunities and increase the number of students who complete a primary education. During the 1980s, concern shifted to simultaneously balancing increased demands for education while continuing to reduce the unit cost of operating an educational system (World Bank, 1981; Psacharopoulos and Woodhall, 1985; Lewin, 1987; Lockheed and Hanushek, 1988; Windham, 1988).

Regardless of what aspect of the argument for internal efficiency is emphasized, educational planners need information about how students progress through the school system (student flow) to assess and improve the efficiency of the system (UNESCO, 1972; Davis, 1966 and 1980; Farrell and Schiefelbein, 1974; Haddad, 1979; Thonstad, 1980). Student promotion, repetition, and dropout rates (flow rates), together with achievement test results, are the two outcomes of the education system most widely used to measure student performance and evaluate the internal efficiency of the system. Achievement tests provide information about how much students have learned; flow rates indicate how students progress through the school system toward completion.

How much students learn and how many students dropout, repeat, or successfully complete a grade are all dependent upon policies that affect which children gain access to school by favoring (often unknowingly) children of a particular social class, ethnic group, gender, etc. These policies include those that determine the amount of direct and indirect resources assigned to education, and the administrative regulations regarding promotion (Schiefelbein, 1980; Lockheed and Hanushek, 1988). Inefficiencies in the educational system can occur because there is insufficient information upon which to design and implement these policies. Therefore, in order to improve internal efficiency, ministries of education need to collect accurate information about student achievement and student flow.

Information on student flow is of particular significance in the formulation of policies aimed at improving internal efficiency because school "wastage" (dropout and repetition) increases the per pupil cost[3], blocks access to eligible school-age children, and prevents students from gaining the basic education that will provide social and economic opportunities later in life (UNESCO, 1972; McGinn and Davis, 1969; Levy, 1971; Dominguez, 1980). Student flow rates are important in forecasting future enrollment, resource allocation, and cost. According to the World Bank, "a small improvement in the efficiency of student flows would free hundreds of millions of dollars for education per year every year." (Dominguez, 1980, p.1). Schiefelbein (1975) has estimated that more than 300 million US dollars are devoted each year to student repetition of the first grade in Latin America. He adds that "just a fraction of this amount is more than all the foreign aid devoted to education in Latin America."(p. 486). In terms of cost per graduate, measured in monetary or non-monetary terms, student repetitions are more costly than dropouts because they reduce the intake capacity of the system, cause overcrowding in classrooms, and, in most cases, reduce the likelihood of eventual graduation (Berstecher, 1970; UNESCO, 1972 and 1977; McGinn and Davis, 1969; Levy, 1971; Dominguez, 1980; Schiefelbein, 1975).

The constraints of austerity and efficiency, within the context of a continuing need to increase access to education, have forced policymakers and administrators in Third World countries to improve planning, change management styles, and evaluate ongoing and future programs on a systematic basis. As part of that process, more resources have been devoted to research and there has been greater reliance on the information provided by research (Myers, 1981). All of these changes have been supported, and sometimes promoted, by international donors and lending agencies.

In the last three decades there has been a sustained effort by these agencies to introduce more sophisticated data monitoring and accounting systems and provide training in this new technology. The introduction of microcomputers, the reduction in their cost, and the development of advanced programming languages (i.e., True Basic, Turbo Pascal, C) and application generators (i.e., LOTUS, DBASE) during the 1980s have provided additional incentives for the introduction of computer-based educational management information systems (Strudwick, 1986; Green, 1986; Cassidy, 1986; Chapman and Boothroyd, 1988). The advocates of this technology argue that the simple and easy-to-use educational databases create the potential for a dramatic expansion in the amount of information policymakers can use to clarify policy issues and evaluate competing policy alternatives (Cassidy, 1986; Evans, 1986; Chapman, 1988). However, more information will not lead to better decisionmaking if the data that support those decisions are not reliable.

Section II: Overview of the Quality of Educational Information in Third World Countries

Reliability of Student Flow Data

Both the introduction of educational management information systems and the successful implementation of new management and planning practices rely heavily on the use of dependable information. Reports from research conducted in many Third World countries (Schiefelbein, 1975, 1979, 1981 and 1985; Cuadra, 1987; Chapman, 1988; Cuadra et al., 1988), as well as in the United States (Oakes, 1986; Plisko, Ginsburg, and Chaikind, 1986; Murnane, 1987), demonstrate that many of the basic educational statistics produced by education offices and used by policymakers are not reliable.

...each school that was visited had a different form for keeping records of students.

An evaluation of the flow of data from schools to the Ministry of Education in Nepal[4] found that the kinds of information collected varied from school to school due, primarily, to the lack of a standard form for data collection. In most of the schools the only information collected at the beginning of the school year was the name of the student; very few schools kept records on the progress of the students, such as frequency of repetitions and promotions. Consequently, when the Ministry of Education requested this information, teachers and headmasters relied upon their memory to respond. This study also found a tendency for schools to inflate the number of student enrollments reported, because the number of teachers assigned to a school is based on the number of students. Data were also often inaccurate simply because there was little awareness at the school level of how important the data were; data collection and reporting were considered time-consuming obligations that offered little benefit to the school. Due to the relative lack of visits by supervisors, information from remote areas was generally less reliable than information from urban centers.

Chapman and Boothroyd (1988) found discrepancies of 25 percent between the enrollment records sent from a sample of 35 primary schools near the capital of the Yemen Arab Republic to the Ministry of Education, and the summary reports produced by the Ministry. They discovered that "data for only 19 of 35 classrooms (54 percent) were correctly transferred" (p. 421) and, for one classroom, "the school report of 141 girls was transposed to 141 boys" (p. 421). A team of researchers from the BRIDGES project visiting rural and urban schools in Yemen found that schools generally keep information that school administrators can use to identify students, such as name, place of residence, and name of the father or mother. Although schools are required to keep records of the student's age and academic status in the grade, promoted or repeating, this is rarely done. In addition, each school that was visited had a different form for keeping records of students. "The lack of a standardized school book-keeping system forced each headmaster to design the school form according to his/her understanding and experience" (Cuadra et al., 1988, p. 152).

In 1987, Chapman and Boothroyd[5] conducted a survey of headmasters and Ministry of Education officials in Botswana and Somalia in order to investigate the quality of enrollment data. They found that, in both countries, Ministry of Education officials estimated that enrollment data have a 17 percent error; in Botswana, headmasters suggested that the data have a 20 percent error and, in Somalia, a 23 percent error. When officials and headmasters in both countries were asked about the accuracy of enrollment data, they rated it as only somewhat accurate. Yet all of the respondents stressed the need to improve data quality.[6]

An evaluation of student repetition data in fourteen Latin American countries, conducted by Schiefelbein (1975), revealed that, with the sole exception of Uruguay, the majority of those countries

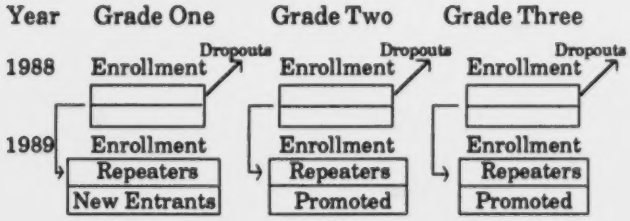
underestimated repetition rates. Schiefelbein states that "published repeater rates generally are one-half the size of the real repeater rate in first grade" (p. 485).

A study conducted by a research team from the BRIDGES Project found that official statistics published by the ministries of education in six Central American countries tended to underestimate, sometimes by more than 100 percent, student repetition rates (Cuadra, 1987). Estimates of the number of student promotions, new entrants, and dropouts in these countries are reached through the grade/retention method (UNESCO, 1972; Schiefelbein and Davis, 1980), which relies heavily on the number of students reported as repeating each grade. Because repeaters' rates are underestimated, these estimates also tend to be biased.

Description of the Grade/Retention Method

According to the grade/retention method of calculating student flow data, enrollment in a grade in any given year is equal to the number of students reported as repeating that grade plus the number of students who are "new" to the grade (those promoted from the previous grade or, in the case of the first grade, new entrants).[7] This method can be represented graphically as follows:

Graph for Grade/Retention Method



When using this method for producing student flow data, schools report only the number of students enrolled in each grade and the number of students repeating each grade to the Ministry of Education. For example, the number of students promoted from first grade in 1988 to second grade in 1989 is equal to the number of students enrolled in second grade in 1989 minus the number of students repeating that grade. The number of students who drop out of first grade between 1988 and 1989 is equal to the number of students enrolled in first grade in 1988 minus the number of students repeating first grade in 1989 minus the number of students promoted to second grade in 1989. This example shows that both the dropouts and the children who are promoted are estimated as a residual between the number of students enrolled in a grade and the number of students repeating that grade.[8] Consequently, the accuracy of both the number of students reported as promoted and those described as having dropped out are dependent upon the accuracy of the enrollment and repetition data supplied to the Ministry of Education for two consecutive years.

Assessment of Student Flow Data In Honduras

The estimates of new entrants to schools in Honduras that are presented in Table One were calculated from official statistics published by the Ministry of Education. Since the numbers of new entrants are based upon enrollment and repetition figures, inaccuracies in the enrollment or repetition data will have an impact on the estimates of new entrants. (See Graph for Grade/Retention Method.) For instance, if the number of repeaters reported for 1979 is 25 percent higher (i.e., 64,000 instead of 51,111) then the estimated number of new entrants will go down by almost 13,000 students (from 153,271 to 140,382). The practical implication of this is that there will be 13,000 less places available for new students than policymakers and planners expect.

Table One: Honduran Enrollment and Repetition Data and Estimates of New Entrants to First Grade

Year	Enrollment	Repeaters	New Entrants
1979	204,382	51,111	153,271
1980	209,522	54,563	154,959
1981	211,975	54,721	157,254
1982	230,383	59,400	170,983
1983	236,619	63,003	173,616
1984	244,643	64,666	179,977

Source: Anuario Estadistico. Ministerio de Educacion, Honduras.

The estimates of new entrants presented in Table One, based on official enrollment and repetition statistics collected from schools by the Honduran Ministry of Education, indicate that in 1979 the population entering school for the first time was 33 percent larger than the total number of six year old children.

A student intake rate (i.e., new entrants as a per-

centage of the six year old population) larger than 100 percent means that there are more students entering the first grade than there are children who are six years old. If the population data are accurate, an intake rate larger than 100 percent could only occur if the new entrants to first grade came from several age groups, combining underage, overage, and six-year-old entrants. If Honduras was engaged in an active campaign to enroll all children in schools, then the intake rate could be larger than 100 percent due to the addition of overage and underage children. However, a country can have an intake rate larger than 100 percent for only so many consecutive years before it runs out of eligible children. As Table Two shows, Honduras has had an intake rate of about 130 percent for at least six consecutive years.

Table Two: Reported Number of Students Entering the First Grade for the First Time, Population of Six Year Olds, and Intake Rate for Honduras 1979-85

Year	New Entrants	Six Year Old Population	Intake Rates
1979	153,271	115,437	133%
1980	154,959	119,314	130%
1981	157,254	123,458	127%
1982	170,983	127,845	134%
1983	173,616	132,308	131%
1984	179,977	136,680	132%

Sources: Anuario Estadistico. Ministerio de Educacion, Honduras. For population data; Proyecciones de Poblacion, 1975-2000. Centro Latinoamericano de Demografia (CELADE) and Consejo Superior de Planificacion (CONSUPLANE), Honduras, 1985.

A simulation model developed to demonstrate this situation[9] shows that intake rates of the magnitude presented for Honduras in Table Two are very difficult to reach. On the basis of the enrollment and population data for Honduras, the model indicates that, in order to have intake rates that are close to 130 percent for more than three years, the participation rates[10] in education for children between seven and ten years old would have to be very close to 100 percent. In addition, the system would have had to increase enrollment of six year olds in the first grade from 28 percent to 85 percent over the six year period. Furthermore, after six years, the system would need to attract over 35 percent of the five year old children in the country.

Educational reality in Honduras during the period under consideration was in fact quite different from that described by the reported statistics. In fact, the participation rate for the seven to ten year old age group between 1978 and 1984 never approached one hundred percent. The participation of six year old children in the educational system grew only by five percent during that period. Although the school system of Honduras has expanded over the last few years, as the data in Table Three shows, the expansion has not been dramatic enough to explain the intake rates reported in Table Two.

Table Three: Participation Rates for the Population of Six to Ten Year Olds in Honduras from 1978 to 1984

Age	1978	1984
6	28%	33%
7	67%	80%
8	76%	84%
9	78%	83%
10	77%	78%

Source: Anuario Estadistico. Ministerio de Educacion, Honduras.

Based on the preceding reasoning, it does not appear that the large number of new first grade entrants reported for Honduras between 1978 and 1984 can be explained by the enrollment of overage and underage students. An alternative explanation for the apparently large intake rates is that the official statistics either underreported the number of students repeating the grade, overreported the total enrollment, or simultaneously did both. This would distort the resulting figures for new entrants since the calculations for new entrants are based on these numbers. (See Graph for Grade/Retention Method.)

In order to test the validity of this hypothesis, a contrasting method for obtaining student flow data must be used to generate an alternative set of statistics. An ideal mechanism for collecting the basic information needed to estimate flow rates is a true pupil accounting system and cumulative record histories (Sammak, 1981). However, these data collection strategies rely heavily on individual record keeping systems, are very expensive to maintain, and, to be successful, require a long, continuous effort by the government. They are, therefore, beyond the current technologies and budgets of most Third World coun-

tries. Until the time that such data collections systems are in operation, planners and policymakers in education need an alternative method of collecting reliable information about how students move and progress through the system in order to evaluate its performance and to formulate policies and strategies for the future.

Section III: An Alternative Approach to Producing Reliable Estimates of Student Flow Data

Description of the Age/Grade Method

One method for obtaining student flow data is the age/grade approach.[11] The age/grade method is based on student enrollment statistics by age and grade for two consecutive years. The basic idea behind this method is that the enrollment by age and grade of any single academic year is related to the enrollment of the following year by a set of specific promotion, repetition, and dropout rates. These rates can be estimated by comparing the enrollment by age and grade for two consecutive years. This is done by using two unrestrictive assumptions, or informed guesses, about student flow. The first assumption concerns the percent of students repeating the last grade (in this case, grade three), and the second the percent of non-promoted students who drop out. A short example will help to clarify this method.

The following table presents the computation matrix needed to estimate the number of promoted students, repeaters, and dropouts for a three grade educational system. The second column has the enrollment by grade of the students who were eight years old in 1988. The last column lists the enrollment by grade for the same age group one year later, when they were nine years old.

The reliability of the student flow rates produced by the age/grade method will depend on the accuracy of the data used, ...

Table Four: Example of a Computation Matrix Used by the Age/Grade Method, in a Three Grade System

Grade	1988 Enroll 8 Yr	Promoted	Non-Promoted	Repeaters	Dropout	1989 Enroll 9 Yrs
1	11,500	8188	2960	2812	148	11,000
2	11,000	8540	2000	1960	40	10,500
3	10,000	9000	n.a.	500	n.a.	9,500

Computation by this method begins by looking at the enrollment in the highest grade (in this example, grade three), during the first of two consecutive years (1988), and making an informed guess about the proportion of students in that grade who will be enrolled as repeaters in the same grade the following year. For example, if we assume that five percent of the students enrolled in grade three repeat, then the total number of nine year olds who are enrolled as repeaters in grade three in 1989 is 500. The repeaters are then subtracted from the 1989 grade three enrollment to get the number of students who were promoted from second grade. Since 9500 minus 500 equals 9000, then 9000 is the number of students who were promoted from grade two in 1988. Next, the number of non-promoted students from grade two in 1988 is calculated as the difference between the enrollment in that grade and the number of students who were promoted to the next grade (grade three) the following year (1989). Finally, the estimated number of non-promoted students is split between repeaters and dropouts by assuming that a fixed proportion of the non-promoted will drop out of school.[12]

In Table Four, we assumed that two percent of the non-promoted students (two percent of 2000), or 40 students, dropped out of second grade. By subtracting the dropouts from the total of non-promoted students (2000 minus 40) we find the number of students repeating grade two, i.e., 1960. This number, the number of students repeating second grade (1960), is then used to calculate the number of students who were promoted from first grade, and the calculations are repeated until the repeaters in first grade are estimated (Schiefelbein and Grossi, 1981).

The reliability of the student flow rates produced by the age/grade method will depend not only on the accuracy of the data used, but also on the ability of the algorithm that the model is using, to reproduce the reality that it represents. To test the age/grade method's algorithm we included it in a simulation model. This model was based on an earlier simulation model designed by Ernesto Schiefelbein, and further developed by BRIDGES to produce repetition, promotion, and dropout rates by grade from data

on enrollment by age and grade for one year on a six grade school system. A second year of enrollment by age and grade was deduced from a set of hypothetical student flow rates. Three different scenarios were tried; they represented a system with high level repetition rates, especially in the first grade, medium level repetition rates, and low level repetition rates. From this data the model produced the three hypothetical sets of transition rates that were used originally to deduce the second year of enrollment by age and grade. These rates were then compared to the original ones used in the three scenarios. In all three cases the age/grade algorithm used in the model accurately reproduced the original flow rates that linked the two consecutive years of enrollment by age and grade data.

Comparison of the Age/Grade Method with the Grade/Retention Method

There are three main differences between the grade/retention and the age/grade methods.

- First, the data required to produce the estimates are different: the grade/retention method requires data on enrollment and repetition by grade, while the age/grade method requires data on enrollment by age and grade.
- Second, the grade/retention method treats repetition as an input, while the age/grade method produces repetition estimates as an output. The grade/retention method uses the repetition figures reported by each school as the key statistic in the estimation of the number of new entrants, promoted students, and dropouts. The age/grade method estimates repetitions, promotions, and dropouts through the use of enrollment figures by age and grade.
- Third, the estimates of the grade/retention method are easier to obtain than the age/grade method's estimates, as the latter are based on a series of iterations and conditional statements that are difficult to handle when computations are done by hand. However, the availability of a microcomputer can facilitate these calculations.

Application of the Age/Grade Method to Estimate Flow Rates in Honduras

A microcomputer simulation model using the age/grade methodology has been utilized to compute a series of repetition, promotion, and dropout figures for up to an eight year period in six Central American countries by both the BRIDGES Project and the World Bank (Schiefelbein et al., 1985; Cuadra, 1987). In Honduras, this microcomputer simulation model yielded estimates of repetition rates for the first grade that were almost twice as large as the figures produced by the standard grade/retention method. The differences between these two sets of data are presented in Table Five.

Table Five: Comparison of Estimates of Repetition Rates in the First Grade for Honduras

Year	Grade/Retention Method (MOE Estimates)	Age/Grade Method (Model Estimates)
1979	27.3%	51.8%
1980	26.7%	54.1%
1981	26.1%	57.6%
1982	28.0%	50.1%
1983	27.3%	50.7%
1984	27.3%	51.7%

Because of the differences between the Ministry of Education repetition figures and the repetition estimates generated by the age/grade model, the estimates of new entrants will also differ substantially between the two approaches, since the number of new entrants is calculated as the difference between enrollment and repetition. The higher the number of repetitions used in these calculations, the lower the estimated number of new entrants will be and, vice versa, the lower the number of repetitions, the higher the estimated number of new entrants.

Table Six presents the estimated number of new entrants that was calculated using repetition rates from the age/grade method in Table Five. A comparison of the estimates from Table Five and Table One indicates that the number of new entrants estimated through the age/grade method is almost twice as large as the estimates obtained using the Ministry of Education data on student repetitions.

Table Six: Comparison of Estimated Number of New Entrants to First Grade To Six Year Old Population

Year	Grade/Retention Data	Age/Grade Estimates	Six Yr Old Population
1979	153,271	97,056	115,437
1980	154,905	97,508	119,314
1981	157,254	95,411	123,458
1982	170,983	120,102	127,845
1983	173,616	117,966	132,308
1984	179,977	119,482	136,680

The estimated number of new entrants to the first grade obtained through the age/grade method is well below the estimate based on the repetition data reported by schools to the Ministry of Education. According to the age/grade estimates of new entrants, the intake rates range from 77 percent to 94 percent of the total six year old population, while the Ministry of Education rates ranged from 127 percent to 134 percent.

Intake rates in the range of 77 to 94[13] percent more closely represent the enrollment patterns of Honduras' education system. The numbers of new entrants that were calculated using the age/grade method can be tested against the same hypothetical scenario as were the higher figures produced by the grade/retention approach. Using the new entrant estimates generated by the age/grade model, the participation rate of the six year old population would grow from 28 percent in 1978 to 85 percent in 1984; the participation rate of the seven year old population needs only to increase from 67 percent to 68 percent and does not need to reach the 100 percent required by the grade/retention estimates. In addition, the estimates of new entrants produced by the age/grade method do not require the participation of five year old students as does the grade/retention method. In fact, the estimates of new entrants generated by the age/grade method made it possible to replicate the historical participation rates of the six year old group of students as they occurred in fact between 1978 and 1984 (displayed in Table Three).[14]

Section IV: Empirical Evaluation of the Age/Grade Method

If the age/grade method is to be accepted as an alternative methodology, it must be shown empirically that this method yields accurate estimates of flow rates. The need for assessment is especially evident in light of the fact that the data used by the age/grade method come from the same school sources as the data on the number of student repetitions used by the grade/retention method.

To prove that the age/grade based simulation model yields more accurate estimates of transition rates than the grade/retention method, it is necessary to obtain direct evidence that information on enrollment by age and grade is more accurate than information on student status. The efficacy of each could be established by comparing the student flow information generated by each method with the actual flow of a cohort of students through a school system. If it could be demonstrated that the flow rate estimates produced by the age/grade based simulation model provide a closer description of student flow rates than the current grade/retention estimates, then the age/grade methodology would be established as a more accurate and hence more useful tool for planning, forecasting, and evaluating policy alternatives.

...student flow rates calculated by the grade/retention method... underestimate the number of student repetitions and overestimate both promotions and dropouts.

Research Objectives and Data Collection Strategy

The age/grade simulation model was developed because the student flow rates calculated by the grade/retention method, used by many ministries of education, underestimate the number of student repetitions and overestimate both promotions (including new entrants) and dropouts. The first objective of the present research was, therefore, to assess whether or not the figures produced by the age/grade model were more accurate than those produced by the Honduran Ministry of Education using the grade/retention method.

The age/grade simulation model uses information on enrollment by age and grade for at least two consecutive years. Therefore, a second objective of this research was to assess the reliability of this data as it is collected in schools, as compared to the reliability of repetition data.

There are several strategies for obtaining information on the actual flow of a student cohort. One is to cross-check different information sources within each school and reconstruct the academic history of each student. Another strategy is to establish independent information about each student's progress through the system from sources close to him/her, such as parents. In both cases, the academic information on each student must be collected for at least a two to three year period.

The task for the research team, consequently, was to reconstruct the academic history of a sample cohort of students. The history was based on data collected from school records and on information given by the parents of the students in the sample. Research by Noel McGinn and Russell Davis (1969) has indicated that parents' reports of their child's academic history are accurate. There are at least three reasons why parents' memories may be more accurate than school records: (1) parents have a strong interest in their children's education; (2) when a child is not at school, she/he is at home, visible to parents; (3) parents have fewer children to look after than teachers do, making it more likely that they, rather than teachers, will remember accurately what the child did in a given year. Based on these research findings, parents' information was selected as the criteria against which information from school records would be judged for accuracy. Where a high level of congruency between parent and school data existed, the school data was accepted as accurate;

where there was a high level of disagreement between parent and school data, the school data was considered inaccurate.

Evidence suggests that the manner in which data are collected by schools affects their accuracy. In order to understand what factors influence the quality of data collected in the schools, the third objective of this research was to investigate data collection practices at the school level. An understanding of the data collection process was expected to provide insight into the factors that have led to inaccuracies in the Ministry of Education's data. Therefore, in addition to interviewing the students' parents and tracing student histories as they were recorded in the school records, the researchers questioned school supervisors, principals, and teachers about the process of reporting data to the Ministry of Education.

Description of the Sample

The sample cohort was drawn from first grade students entering school for the first time in 1983. An ad hoc committee, composed of a former deputy from the Ministry of Education, the Director of the In-Service Teachers' Training Unit of the Ministry of Education, the Director of the Teachers' College (Escuela Superior del Profesorado), and the BRIDGES consultant, selected geographical regions which were considered by the Ministry to be representative of "traditional" and "modern" schools in Honduras. The designations "traditional" and "modern" were made by Ministry of Education officials and based upon the economy of the area in which the school was located. "Traditional" regions were those in which primarily subsistence farming took place; in "modern" regions there was some industry, and agriculture was generally market oriented. The regions selected for data collection were Catacamas, a traditional area in the southwest, and El Progreso, a more modern area in the northeast.

The research team, on advice from the Ministry staff, selected urban and rural schools in both regions. Three criteria were used in this selection: (a) a school must have had at least 30 students enrolled in the first grade in 1983; (b) school records from 1983 –1986 must have recorded the names of repeating students; and (c) school records must have been available for the entire 1983–1986 period.

The sample included all students listed as "new" entrants to the first grade in 1983. Students enrolled in grade one in 1982 who repeated the first grade in 1983 were excluded from the sample. In addition, students for whom no current address could be located were also excluded. A maximum sample size for each school was set at 35 students. When the pool of eligible students in the first grade class of a school exceeded this number, students were randomly selected from the enrollment list.

Table Seven lists the enrollment figures for grade one, the estimated number of new entrants in grade one, the number of students included in the sample, and the sample size as a percentage of new entrants for the 13 schools included in the study.

Table Seven: Schools, Grade One Enrollments, New Entrants, and Sample by Region, Location, and Size

Region	School	Grade 1 Enrollment	New Entrants[1]	Sample
Catacamas (southwest):				
Urban	1	276	155	35
	2	69	39	33
Rural	1	110	62	35
	2	32	23	23
	3	39	36	36
	4	34	19	12
	Total	560	333	174
Progreso (northeast):				
Urban	1	243	136	32
	2	213	119	28
Rural	1	120	67	30
	2	147	82	27
	3	61	34	8
	4	65	36	14
	5	31	17	14
	Total	880	493	153
Total		1440	826	327

Note: "New Entrants" are estimated on the basis of parent reported repetition rates for grade one substantiated by an in depth review of grade one records in two schools during sample selection.

Data Collection

Eight interviewers from the Training Unit of the Ministry of Education joined the research team to carry out the field work. Each of them had been a primary school teacher and had worked as a teacher trainer in the Training Unit for at least four years. Some of the interviewers had been principals and supervisors; some had received, or had completed

Parents were visited in their homes and asked to recall their child's grade level, attendance pattern, age, and grade repetitions...

course work towards, a master's degree in educational administration. The interviewers' experience and professionalism were critical factors in gaining the cooperation of principals, teachers, and parents. The interviewers developed interview schedules and designed a data collection form for individual student enrollment and status data.

Three types of data were collected. The first consisted of duplicating the student enrollment and repetition figures that had been sent to the Ministry of Education for the grade/retention calculation of student flow rates. The second type of data was drawn from the same source, the school records, but focused on cross-checking the status of each student in the yearly enrollment lists. For example, if a student was enrolled in the first grade in both 1983 and 1984, his/her status in 1984 was that of repeater, even if the student had not been included in the list of repeaters reported to the Ministry of Education. The third type of data collected was drawn from interviews conducted with the parents of the students in the cohort. Parents were visited in their homes and asked to recall their child's grade level, attendance pattern, age, and grade repetitions from 1984 through 1986.

Contact with students was made by visiting the School District Supervisors responsible for the schools in the study. They were informed of the objectives of the research and asked for their assistance. In general the supervisors cooperated fully and, in some cases, they introduced the research team to the school principals. Home addresses of the students included in the study were usually obtained from the students themselves or from classmates.

The school records were not always well organized or well kept. For example, there were some students whose names differed between their enrollment and their final evaluations, or changed from one year to the next. In these cases the teachers or students were asked to determine a "real" or consistent name. There were no records for some students. Others appeared in the enrollment records, but not in the final evaluation records. In contrast, when parents were asked about the age of children who were attending school, almost all of them were able to produce their child's birth certificate for the interviewer. School information and parental interviews were completed for 327 students in the 1983 sample cohort from grade one.

Data Analysis

Validity of Data on Age and Grade

The comparison between the information obtained from the school records and from parent interviews with respect to the child's age and grade attended year by year shows a relatively high level of agreement. The 1984–86 data on child's age correspond for 95% of the time. The 1984 data on grade attended that were gathered from parents, when compared to data from school records, coincided for 86 percent of the students; 1985 data corresponded for 90 percent of the students; and 1986 data on grade attended was the same for 89 percent of the students. However, if the three years are taken as a total unit, the information did not coincide for 25 percent of the students. The following cases illustrate some of the kinds of discrepancies that occurred.

The first is the case of a student who was promoted from the first grade to the second grade in 1983, but, the following year, was shown as a repeater in the first grade. However, his name also appeared in the 1984 evaluation records for the second grade as being promoted to the third grade and, in 1985, as enrolled in the third grade. The student's parents reported that he had not repeated the first grade.

The second case involves a student whom the school records showed as promoted from the first grade to the second grade in 1983. In 1984 the student was listed as enrolled in the second grade and promoted to the third grade at the end of the academic year. In 1985, however, she was shown as enrolled in the first grade again. At the end of that year, the student was shown as promoted to the second grade. In 1986, she was listed as enrolled in the second grade with the status of a "new" student in that grade, although the 1984 records had also shown her in the second grade. The parents of this student said that she repeated the first grade in 1984. She was then promoted to the second grade in 1985, failed that grade, and then enrolled, as a repeater, in the second grade in 1986. In this case, the parent's and the school record's information only corresponded at the point of the student's original enrollment in 1983.

A somewhat similar situation occurred in the case of a student who was promoted in 1983 from the first to the second grade, and was shown as enrolled in the second grade in 1984. In that year she was listed as having failed the second grade and, therefore, was enrolled in that grade again in 1985. In 1985, however, her name did not appear in the evaluation records for the second grade but, instead, in the

records for the first grade where she was shown as promoted from the first to the second grade. In 1986, she appeared in the enrollment records for the second grade as a student "new" to that grade, although the records had listed her as in the second grade in 1984. Her parents said that the student was promoted from the first to the second grade in 1983 and was subsequently enrolled in the second grade in 1984. However, at the beginning of 1985, they had asked to have her enrolled in the first grade again as they believed that "she was learning nothing." The girl attended first grade in 1985 and was promoted to the second grade at the end of that year. In 1986, according to the student's parents, she was enrolled in the second grade.

These cases illustrate how the high level of agreement in grade attended between the year-to-year information provided by the schools and by the parents does not necessarily indicate that school records are accurate. The likelihood of inaccuracies in reporting a student's status sometime during her/his academic life is greater than the year-to-year figures indicate. If the academic histories of the whole cohort are taken as a single unit over time, instead of being broken into year-to-year segments, then longitudinal comparisons provide a much lower level of correspondence.

Validity of Flow Data

In order to assess the accuracy of the school reports in terms of enrollments, dropouts, repetitions, and the ages of students, direct comparisons have been made among the various types of data collected. For example, information on the number of students repeating each grade has been recorded according to three separate methodologies: (a) grade/retention–based on the number of students reported in the school records as repeating each grade from 1984 through 1986; (b) record comparison–based on tracking students by name through the enrollment records for each grade from 1984 through 1986; (c) parents–based on parent recollections of child's grade level, attendance pattern, age, and grade repetition from 1984 through 1986.

Table Eight presents three reconstructions of the academic history of the cohort based on the three methodologies. As 1983 is the beginning of the period being studied, there is no disagreement among the three reconstructions during that year. The reports of enrollment for grade two in 1984, grade three in 1985, and grade four in 1986 are reasonably congruent across all reconstructions. These figures represent the new students entering each of these grades. However, beginning in 1984, there is increasing disagreement between the grade/retention reconstruction and the parent criterion in terms of repeaters and dropouts. The grade/retention method consistently underreports the number of students repeating a grade level and, as a consequence, over-estimates the number of dropouts. For example, in 1984, the grade/retention method underreports the number of students repeating the first grade by 81 and reports two more students in the second grade than the parents did. A small number of the underreported repeaters, in this instance two students, are reported as new entrants in the second grade. The record comparison method reflects a similar pattern of underreporting of repeaters, with the exception of grade two, but the differences from parent criterion are smaller.

Table Eight: Cohort Academic History as Reconstructed by Grade/Retention, Record Comparison, and Parent Information

Grade/Retention Reconstruction of Academic History

Year	Grade 1	Grade 2	Grade 3	Grade 4	Dropout	Total
1983	327	0	0	0	0	327
1984	63	172	0	0	92	327
1985	17	72	118	0	12	327
1986	4	27	62	82	152	327

Record Comparison Reconstruction of Academic History

Year	Grade 1	Grade 2	Grade 3	Grade 4	Dropout	Total
1983	327	0	0	0	0	327
1984	128	178	0	0	21	327
1985	41	126	126	0	34	327
1986	16	62	111	89	49	327

Parent Reconstruction of Academic History

Year	Grade 1	Grade 2	Grade 3	Grade 4	Dropout	Total
1983	327	0	0	0	0	327
1984	144	170	0	0	13	327
1985	48	121	131	0	27	327
1986	23	56	116	97	35	327

Underreporting the repetition rates in each grade has a cumulative effect when it is analyzed across grades.

The proportion of the cohort reported as progressing a grade level each year, specifically grade two in 1984, grade three in 1985, and grade four in 1986, is fairly consistent. For these specific grade levels in these years the grade/retention basis reports a 52.6, 36.1, and 25.1 percent promotion rate respectively, compared to the 54.4, 38.5, and 27.2 percent reported by the record comparison basis and the 52.0, 40.1, and 29.7 percent reported by parents.

Table Nine presents the differences between the two reconstructions of the academic history of the cohort based on school records and parent reconstruction. This comparison permits a more discrete analysis of the congruence between the two reconstruction methods.

Table Nine: Differences Between Grade/Retention and School Record-Based Reconstructions as Compared to Parent Information

Note: A (-) number is less than and a (+) number is greater than PARENT criterion.

Grade/Retention Reconstruction

Year	Grade 1	Grade 2	Grade 3	Grade 4	Dropout
1984	-81	2	—	—	79
1985	-31	-49	-13	—	93
1986	-19	-29	-54	-15	117

Record Comparison Reconstruction

Year	Grade 1	Grade 2	Grade 3	Grade 4	Dropout
1984	-16	8	—	—	8
1985	-7	5	-5	—	7
1986	-7	6	-5	-8	14

Table Ten summarizes an analysis of the percent of each grade level enrollment that repeated the grade in the subsequent year based on grade/retention, record comparison, and parent information. These percentages cannot be directly derived from Table Eight, with the exception of 1984, due to repeaters from previous years progressing to the next grade level as new entrants. For example, the 72 students reported in grade two in 1985 in Table Eight under the grade/retention academic history, are composed of both repeaters from the 172 students who were in grade two in 1984 and students who had been promoted to the second grade out of the 63 students who had been repeating grade one in 1984. The grade/retention basis of reconstructing the academic history of the cohort substantially underestimates the repetition rates, particularly in the first grade, in comparison to parent and enrollment based estimates.

Table Ten: Percent of Repeaters to Total Enrollment

Grade/Retention Reconstruction of Academic History

Year	Grade 1	Grade 2	Grade 3
1984	19.27%	—	—
1985	26.98%	16.86%	—
1986	23.53%	27.78%	16.95%

Record Comparison Reconstruction of Academic History

Year	Grade 1	Grade 2	Grade 3
1984	39.14%	—	—
1985	32.03%	24.72%	—
1986	39.02%	25.40%	23.02%

Parent Reconstruction of Academic History

Year	Grade 1	Grade 2	Grade 3
1984	44.04%	—	—
1985	33.33%	20.00%	—
1986	47.92%	21.49%	22.90%

Underreporting the repetition rates in each grade has a cumulative effect when it is analyzed across grades. Table Eleven presents the total number of students in the sample who repeated the grade according to each of the three methodologies. The data include all students, regardless of grade level, who repeated a grade in 1984, 1985, or 1986.

Table Eleven: Total Sample Repeaters by Year 1983 to 1986

	Grade/Retention Figures		Record Comparison		Parents	
Year	Number	%	Number	%	Number	%
1984	63	19.3%	128	39.1%	144	44.0%
1985	46	14.1%	85	26.0%	82	25.1%
1986	44	13.5%	74	22.6%	79	24.2%

The grade/retention figures consistently underreport the number of repeaters in comparison to the

The consequence of this underestimation of repetitions is the overestimation of dropouts...

parents' criterion. In 1984, the number of repeaters calculated according to the grade/retention system represents only 43.8 percent of those reported by parents. In 1985, it represents 56.1 percent, and in 1986, 55.7 percent. In contrast, the record comparison method represents 88.9 percent in 1984, 103.7 percent in 1985, and 93.7 percent in 1986.

Because the grade/retention figures are based on the number of student repetitions reported in school records, it is clear that these records consistently report only about half of the actual number of students repeating. Because the school records were also the basis for the record comparison figures, the reported enrollment in each grade by student name appears to be substantially more accurate than the number of repetitions reported.

The consequence of this underestimation of repetitions is the overestimation of dropouts, as seen in Tables Eight and Nine and summarized in Table Twelve. The grade/retention method assumes that, for any given year, all students who are not recorded as repeating are new entrants. For example, for grade one, the number of new entrants equals grade one enrollment minus repeaters. Following this process, dropouts are calculated in the following manner: dropouts in 1984 equal grade one enrollment in 1983, minus grade one repeaters in 1984, minus new entrants in grade two. The underreported repeaters in grade one were counted as both dropouts and new entrants in that grade in 1984. Unless the underreporting of repeaters is compensated for by an overreporting of new entrants in grade two, overestimations of both dropouts and new entrants to grade one are the direct result of underreporting repeaters in grade one in 1984.

There is a high level of congruency between the record comparisons and parents' histories, particularly in comparison to the grade/retention estimates of dropouts. While the grade/retention reconstruction estimated the cohort dropout for 1984, 1985, and 1986 as 28.1, 36.7, and 46.5 percent respectively, the record comparison reconstruction showed 6.5, 10.4, and 15.0 percent, and the parents' histories recorded 4.0, 8.3, and 10.7 respectively. Using the grade/retention based estimates, dropouts in the cohort would be significantly overestimated. Table Twelve summarizes the dropout data for the three reconstruction methodologies.

Table Twelve: Estimated and Reported Dropouts 1983 to 1986

	Grade/Retention Figures		Record Comparison		Parents	
Year	Number	%	Number	%	Number	%
1983	0	0.0%	0	0.0%	0	0.0%
1984	92	28.1%	21	6.4%	13	4.0%
1985	120	36.7%	34	10.4%	27	8.3%
1986	152	46.5%	49	15.0%	35	10.7%

While the grade/retention method estimates the percentage of dropouts for 1984, 1985, and 1986 at 28.1, 36.7, and 46.5 percent respectively, the record comparison reconstruction indicates 6.4, 10.4, and 15.0 percent, and the parents' reconstruction reports 4.0, 8.3, and 10.7 percent respectively. The grade/retention methodology significantly overestimates the number of dropouts in comparison to either the record comparison or parents' information. As stated before, the grade/retention overestimation of dropouts is the direct result of underreporting repetitions, which also leads to overestimating the number of new entrants.

Two final tests of correspondence between school data and parent data were conducted. In 1983 parental and school records were in full agreement on the students' grade level, in 1984, 86.2 percent, in 1985, 90.5 percent, and in 1986, 97.2 percent. This is an average correspondence between school grade enrollment records and parental reports of 93.5 percent per year. School records and parents were in agreement on the specific ages of the students in 93.3 percent of the cases in 1983, 94.2 percent in 1984, 93.3 percent in 1985, and 94.5 percent in 1986. Although this is a high mean correspondence (93.9 percent per year), due to the fact that parents are legally required to present the birth certificate of the child on his/her first enrollment, a higher level of agreement might have been expected. However, both of these records are much more accurate than the reported number of repeaters. As a consequence, a projection methodology based on school records of age and grade will be more accurate than a methodology based on the reported number of repeating students.

Conclusion

The hypothesis underlying this research was that when parent and school information have high levels of congruency, then the school information is

likely to be as valid as particular information and useful for planning and projection purposes. Based on a comparison of school records of repeaters, school records of grade enrollment by name, and parent-supplied information, the following conclusions have been drawn:

(a) The school reports of grade level enrollment are valid based on a high level of congruency between parent and school information;

(b) The school reports of student age at each grade level are valid based on congruency between parent and school information;

(c) The school reports of repeaters by grade are not a valid reflection of the actual situation based on a high level of disagreement between reported repetition status and school enrollment and parent information.

The implications of these conclusions are:

The information that schools send to the Ministry of Education regarding the grade attended by students and their ages is fairly accurate. Therefore, the Ministry data base includes valid information for the use of the age/grade approach to estimating national student flows in Honduras.

The reports of the number of repeating students that schools send to the Honduran Ministry of Education are not valid. Therefore, the current Ministry of Education projection method will underestimate repeaters and over estimate dropouts and new students to grade level.

Section V: Factors Affecting Data Accuracy

One of the objectives for this research was to gain a better understanding of the types and quality of data the Honduran Ministry of Education collected. Based on information from interviews with teachers and school directors, and on an examination of the enrollment records in the schools, a picture of the reporting process emerged. The enrollment report that teachers complete at the beginning of the school year in March contains the individual student's name, birth date, address, and whether the student is a repeater. A copy of this report is sent to the School District Supervisor and the original is maintained in the school's files. Approximately a month after school begins, the Ministry of Education sends the schools the Initial Consolidated Enrollment Report. This report requests data on age and number of repeating students by grade level. Official procedure stipulates that reporting should be based on the school enrollment record. Toward the middle of the school year in June/July, the Final Consolidated Enrollment Report is completed by teachers and submitted to the Ministry through the area supervisors, by school principals. At the end of the school year, in November/December, student evaluations, which report promotions, failures, and names of students not evaluated, are completed by the classroom teacher and submitted to the Ministry by school principals through the respective district and provincial supervisors.

Teachers and principals were interviewed about how they collect enrollment data, how they record the data, and how they fill out the Initial and Final Consolidated Enrollment Reports. From these interviews it became clear that only the age of the student is cross-checked against an official document when the data are being collected and transferred to the school enrollment books. For students' status (repeater or new to the grade) information, teachers rarely request the evaluation certificate from the previous year and generally rely upon their own memories or on information provided by the student. It is not unusual for the Initial and Final Consolidated Enrollment Reports to be completed by teachers during class hours, frequently by simply asking all of the repeaters in the class to stand up. Information about the student's experience in the previous school year comes, in almost all cases, from the students. When teachers are asked to fill out reports on the number of students enrolled in their classes and on the status of the students, they usually ask the student to report on her/his status. Students who leave the school before the end of the academic term may be counted as new when they return to school the following academic year.

The educational system in Honduras may include incentives which lead teachers to underreport student repetitions. For example, Article 144 of the law that created the System of Promotion and Evaluation (Sistema de Evaluacion y Promocion Cortrolada) states that, if more than 15 percent of the students fail a grade, then the teacher will be held accountable for the failures and the principal and supervisor will study the reason for the teacher's incompetence in order to overcome his or her limitations.[15]

The interviews with parents provided examples of how teachers try to avoid appearing to have failed too many children.

The interviews with parents provided examples of how teachers try to avoid appearing to have failed too many children. In some cases the teachers do not give the parents the evaluation certificate at the end of the year, thereby forcing them to enroll their child in the same grade again the next year. Some teachers asked parents to enroll a student in the grade he or she had just completed, although the child had passed the grade. The argument given to justify this request was that, although the student had been promoted, he or she had not fully mastered the basic

skills required to be successful in the following grade and, therefore, it was in the student's interest to repeat the grade. There were also cases in which parents demanded that their children repeat a grade because they felt that the child "had learned nothing" or "did not know how to read," in spite of the teacher promoting the student.

The age/grade method for estimating student flow does not depend on information provided by students. Students' ages are obtained from birth certificates, which are the screening device that schools use to select new entrants to the first grade in the current crowded conditions. This helps to explain why the interviewers encountered no problems when they asked parents to show them the birth certificates for their children. Generally these certificates were kept in a safe place in the house; in most cases they showed few signs of deterioration. Parents said that they had been required to bring this certificate to registration at the beginning of the school year. The care with which the parents protected these certificates is related to both the time it takes to obtain one from the government and the essential role it plays in school registration.

Section VI: Recommendations for Improving Educational Data

The need for austerity measures has dominated the discussion of priorities and policies in many Third World countries due to a sustained demand for educational opportunities, declining economic resources, and a reduction in the rate of growth of educational expenditures. These conditions have forced educational planners, policymakers, administrators, and donor agencies to renew their commitment to increasing the internal efficiency of educational systems. The process of identifying appropriate policies to help reduce internal inefficiencies in education is facilitated if policymakers and planners have sufficient, accurate information upon which to base their actions and recommendations.

An assessment of the quality of the data used by ministries of education in several Third World countries indicates that the statistics used to measure several indicators of internal efficiency, such as repetition figures and estimates of flow rates, are neither reliable nor valid. Although educational planners, policymakers, administrators, supervisors, and teachers are, in general, aware of the lack of accuracy in educational statistics and agree upon the need to have more reliable statistics, little is being done to improve data collection or produce more accurate data. One reason for this is that the people working in education are not aware of alternative strategies for calculating more reliable estimates of internal efficiency. Another reason is the lack of information about the causes of low validity.

One partial solution to the problem of low data quality is to design data collection procedures and strategies that reduce mistakes in data collection and reporting. One strategy is to instigate pupil accounting systems and cumulative record histories. This strategy, however, is expensive and demands computer technologies that are sophisticated relative to current budgets and technical expertise in most Third World countries.

This study provides only a preliminary assessment of the repetition, grade level enrollment, and age information in the Honduran educational data base, but it strongly supports the use of an age/grade approach to calculating new entrants to grades, repeaters, and dropouts. The inherent inaccuracy of the repetition reports and the apparent incentive to underreport the actual number of repetitions in each grade suggest that projection methodologies relying on such data will produce inaccurate pictures of student flows. In countries in which the ministries of education collect enrollment data by age and grade, the problem of inaccurate estimates of student flow rates can be overcome by using an age/grade methodology, providing that the age information reported by schools is accurate. As age is widely used as a screening device to admit children to the first grade, accurate age information can be guaranteed by the administrative requirement that parents present a birth certificate when they enroll their children at the beginning of the school year.

The major factor that might make ministries of education hesitate to use the age/grade methodology is the amount of calculations required. However, the use of microcomputer-based simulation models, such as the one developed under the BRIDGES Project, can, to a large extent, solve this problem.

In addition to adopting an age/grade approach to generating student flow data, ministries of education in Third World countries can improve their present systems for the collection of educational data. At least three types of obstacles currently threaten data integrity: (1) incentives that promote inaccurate reporting, such as teachers' evaluations being dependent upon student achievements, or school resource allocations being tied to enrollment quotas; (2) a lack of clarity in definitions, procedures for data collection, and instruction for data reporting, as well as a lack of standardized data collection instruments; (3) a failure by teachers and headmasters to understand why accurate student flow data are important.

...a system of incentives for teachers and headmasters that promotes a more active role for them in the data collection process could be introduced.

In countries where there are disincentives to accurate reporting, those disincentives should be removed. If they cannot easily be removed, then it is necessary for the government to implement a system that monitors more closely the process of data collection and reporting. In addition, counter incentives, which reward the collection and reporting of accurate data, such as providing opportunities for participation in training seminars, could be introduced. One method by which to enhance monitoring would be to train supervisors in basic data analysis techniques that will help them check for consistency.

In many countries the quality of the information reported by schools can be improved by simply reducing the load of data that the ministry of education demands. An inexpensive and easy to implement initiative that can contribute enormously to an improvement in the quality of information collected by schools is the use of standardized data collection forms. These forms should be made available each year to all schools before the enrollment period begins. Schools should also be provided with standardized record keeping books to make school data more homogenous. Such tools would also reduce the time required for training teachers in data collection techniques. These forms might serve as an incentive for teachers and headmasters to pay more attention to the process of data collection, as they indicate the commitment of the ministry of education to accurate statistics.

Teachers and headmasters frequently have not been informed of the purposes of data collection, nor has the importance of their role as data collectors been made explicit to them. This leads to their misunderstanding how enrollment data affect school management and resource allocation. This misunderstanding is reinforced by the failure of the ministry of education to disseminate back to the schools and districts enrollment information that they have collected from them. An easy to implement strategy for overcoming this problem would be for the ministry of education to send each school a summary report of the educational statistics for its district, and a comparison of these statistics with district goals. District supervisors could receive a more detailed report. In addition, a system of incentives for teachers and headmasters that promotes a more active role for them in the data collection process could be introduced.

Endnotes

1. A "gross enrollment ratio" is the percentage relationship between total enrollment and the estimated school age population for each educational level.
2. Articles by Schultz (1961) and Becker (1964) are frequently quoted as the most influential formal presentation of this argument, which has been labeled the "human capital theory of economic development".
3. Often in the range of 30 to 100 percent in Third World countries (Haddad 1979).
4. This evaluation was conducted by Dr. Madhup Dhungana, Director of New ERA, a nonprofit organization that was commissioned by the government of Nepal to assess the quality of data being used by the Ministry of Education. Dr. Dhungana was interviewed by Ernesto Cuadra (Cuadra et al., 1988).
5. Results of this survey were included in a preliminary version of the article by Chapman and Boothroyd published in 1988.
6. When headmasters were asked to judge the accuracy of the data on a scale from one to four, where one represents extremely accurate and four represents a total lack of accuracy, the average rating of headmasters from Botswana was 2.62 and, of headmasters from Somalia, 2.64. The mean rating by Ministry of Education officials from Botswana was 2.88 and, by officials from Somalia, 2.93. When the same headmasters and officials were asked "to what extent do you believe that improving the quality of education data is important, given other priorities of the Ministry of Education?" the headmasters of Botswana responded with an average rating of 1.49 and the officials with a rating of 1.56, one indicating great importance and four little significance. The average rating given by the headmasters from Somalia was 1.21 and 1.07 by the officials.
7. This relationship is described by the following formula: $E_{t,g} = N^{n}_{t,g} + R_{t,g}$ where: $E_{t,g}$ is the number of students enrolled in grade "g" in year "t". N^{n} is the number of "new" entrants to grade "g" in year "t", or the promoted from grade "g-1" in year "t-1". $R_{t,g}$ is the number of students enrolled as repeaters in grade "g" in year "t".
8. This relationship is described by the following formula: $D_{t,1,g} = E_{t,1,g} - N_{t,g} - R_{t,g}$ where: $D_{t,1,g}$ is the number of students who dropout from grade "g" between year "t-1" (previous year) and year "t" (current year). $E_{t,1,g}$ is the number of students enrolled in grade "g" in year "t-1"(previous year). N^{n} is the number of "new" entrants to grade "g" in year "t", or the promoted from grade "g-1" in year "t-1". $R_{t,g}$ is the number of students enrolled as repeaters in grade "g" in year "t".
9. The results and methodology for this simulation are presented in Appendix One. For example, the model adopted the very optimistic assumption that all of the children who were six to ten years old in 1978 would have attended first grade by 1981. Based on that assumption, after 1981, the only over-age students would be the seven year olds. In order to gradually reduce the number of seven year old students enrolled in the first grade, the model assumed that, between 1978 and 1984, the system increased participation in the first grade of six year old children from 28 percent to 85 percent. According to the model, the proportion of six year old children entering the first grade between 1979 and 1984 would be 31%, 36%, 43%, 54%, 66%, and 85%. The prediction of these two assumptions would be that every year the system needs to attract an increasing proportion of five year olds in order to reach the intake rates of Table Two. If the intake rates continued to be over 100 percent, then the system would need to attract four year old children in increasing numbers.
10. The proportion of students of an age group who are attending school.
11. This method is described by Schiefelbein and Grossi in their 1981 "Statistical Report on Repetition in Latin America" (pp. 26-42).
12. In 1986 a BRIDGES Project research team carried out more than thirty experimental applications of this method to estimate repetition in six Central American countries. Their work showed that the final estimates of the number of repeaters are not very sensitive to (i.e., dependant upon) informed guesses about the proportion of dropouts or the proportion of students who repeat the highest grade (Cuadra, 1987).
13. According to the age/grade estimates of new entrants the intake rate for 1979 is 84%; for 1980 is 82%; for 1981 is 77%; for 1982 is 94%; for 1983 is 89%; and for 1984 is 87%.
14. See Appendix for a more detailed presentation of these results.
15. The text in Spanish says: "Cuando la reprobacion escolar exceda del 15% de los alumnos evaluados se imputara, en parte, la responsabilidad a la capacidad docente del maestro, en este caso el Director de la Escuela y el Supervisor Escolar estudiaran las causas con el proposito de superar las limitaciones del maestro".

Appendix

In Third World countries it is not uncommon to find student intake rates for first grade that are larger than 100 percent. Since an intake rate is estimated by dividing the number of new students entering first grade by the entrance age population, the standard explanation for intake rates larger than 100 percent is that new entrants to first grade come not only from the entrance age group but from other age groups as well. This explanation can only justify the occurrence of intake rates larger than 100 percent for a period of three to four years; beyond that time period the pool of potential new entrants from ages seven and above becomes too small.

The simulation model presented in this appendix was developed to show that under normal circumstances it is highly unlikely that countries will have intake rates higher than 100 percent for more than six years in a row. All the examples presented here used data from Honduras for the 1978 - 1984 period.

The model has four interrelated parts. The first one contains the basic information for estimating what proportion of the total population of six to ten year olds have never attended school, i.e. the pool of potential new entrants. The second part contains a series of enrollment rates. These rates are used to simulate the proportion of each age group that is enrolled in first grade for the first time each year. The third part consists of population data for the five and six year old group; this data is used to get the baseline data needed to run the model. The fourth part of the model uses the information from part one, two and three to estimate how many new entrants to first grade come from each age group.

The estimate of the size of the pool of potential new entrants to first grade (the first part of the model) is done in two steps. First, the participation rates for the six to ten year old population are calculated by dividing the enrollment of each age group by its population. Then an assumption is made regarding what proportion of the six to ten year old group attended school before the base year for the model and dropped out. The size of the pool of new entrants is then estimated by subtracting from the population of each age group both the number of students who are enrolled in the base year, and the number of students who left the system before the base year.

For example, in Honduras the participation rate of the six year old group in 1978 (the base year for this example) was 28 percent. This leaves 72 percent of the population of this age group who can be considered as potential new entrants into first grade in 1979. This is equal to 80,435 children. In building the model we assumed that there were no five year olds in the system prior to 1978, therefore the dropout rate before 1978 for the six year old group is zero. For the 10 year old group, the participation rate in 1978 was 77 percent. This means that 23 percent of this age group can be considered potential newcomers to first grade in 1979. However, since it is likely that a fraction of this 23 percent might have attended school before 1978 and dropped out, the pool of potential newcomers from this age group must be reduced by an amount equal to the number of students who dropped out before 1978. In the example shown on Table A1 it is assumed that 15 percent of the population of this age group was enrolled prior to 1978 and dropped out. This means that the pool of potential newcomers to first grade from the ten year old group is only equal to 8 percent (23 minus 15 percent) which is equivalent to 7,678 children.

Table A1: Population and Enrollment by Age
Honduras 1978

Age	Enroll in all grades 1978	Population 1978	Participation rates 1978	% Dropout Before 1978	Pool of Potential New Entrants
6	31,281	111,716	28%	0%	80,435
7	71,919	107,341	67%	2%	34,349
8	78,476	103,258	76%	6%	19,619
9	77,549	99,422	78%	11%	11,931
10	73,798	95,842	77%	15%	7,678

The second part of the model consists of Table A2 which contains enrollment rates in first grade for the five to eleven year old age groups. These rates show the proportion of each age group which is enrolled for the first time in first grade each year. For example,

column two of Table A2 indicates that in 1979 31.8 percent of the six year old Honduran population were presumed to have been enrolled in first grade for the first time. This column also shows that 60 percent of the seven year olds, 21 percent of the eight year olds, 12 percent of the nine year olds, and 7 percent of the ten year olds and 8 percent of the eleven year olds were assumed to have attended school for the first time in 1979.

Table A2: Hypothetical Enrollment Schedule Into First Grade for Ages Six to Eleven in Honduras: 1979 -1984

Age	1979	1980	1981	1982	1983	1984
5	0.0%	0.0%	0.0%	0.0%	29.6%	30.4%
6	31.8%	43.4%	64.7%	99.6%	100.0%	100.0%
7	60.0%	59.7%	56.6%	35.3%	0.4%	0.0%
8	20.7%	12.0%	8.5%	0.0%	0.0%	0.0%
9	12.0%	10.3%	0.0%	0.0%	0.0%	0.0%
10	7.3%	6.0%	0.0%	0.0%	0.0%	0.0%
11	8.0%	3.7%	0.0%	0.0%	0.0%	0.0%

Note: In 1981 all of the children that were either 6, 7, 8, 9 or 10 years old in 1978 would have attended first grade. From 1981 on there are no children over 8 years old who have not attended first grade.

The rates on Table A2 are defined by the model. They are based on the participation and dropout rates presented in Table A1 and on the assumption that all members of a cohort attended first grade at least once before they were twelve years of age.

Because the model works on the assumption that all the members of a cohort will attend first grade before they are twelve years old, starting with the six year old group in 1979, the addition of the percents displayed on the diagonal of Table A2 with the corresponding participation and dropout rates presented in Table A1 will add up to 100 percent. For example, according to the figures in Table A2 it is assumed that 31.8 percent of the six year old group came to first grade for the first time in 1979; in 1980, when this age group was seven years old, it is assumed that 59.7 percent were enrolled for the first time in first grade, and in 1981, when the children in the cohort were eight years old, it is assumed that 8.5 percent of them were enrolled for the first time in first grade. By 1981, therefore, 100 percent of the cohort is assumed to have been enrolled in first grade. For the age group that was nine years old in 1979 the 100 percent enrollment in first grade is figured by adding the 76 percent participation rate from 1978, when the children in the cohort were eight years old (Table A1), to the six percent drop out before 1978 (also from Table A1), to the 12 percent enrollment in first grade in 1979, to the 6 percent enrollment in 1980 when the children in the cohort were 10 years old.

The third part of the model only contains population data. This data is used as the base line to estimate the total pool of potential new entrants to first grade. This pool is then distributed into Table A4 using the rates estimated on Table A2.

Table A3: Population of Five and Six Year Old

Age	1979	1980	1981	1982	1983	1984
5	120,142	124,177	128,490	133,056	137,701	142,251
6	115,437	119,314	123,458	127,845	132,308	136,680

The fourth part of the model uses the figures on new entrants that are computed with the official enrollment and repetition data from the Ministry of Education, the population data from Table A1 and A3, the participation and dropout rates from Table A1, and the hypothetical participation rates used in Table A2 to estimate the number of students enrolled for the first time in first grade each year. The model starts by estimating the number of seven year olds enrolled in first grade. This enrollment is calculated using the population data of Table A1 and the participation rates of Table A2. Then the model estimates the first grade enrollment of the eight, nine, ten, and eleven year old groups. Once this is done the model estimates the first grade enrollment of the six year old group. The enrollment from this age group is computed as a residual between the number of new entrants and the number of students ages seven to eleven who come for the first time to first grade each year. Finally the model estimates the enrollment to first grade that comes from the five year old group. As for the five year old age group, this enrollment is also computed as a residual. This time it is the residual between the number of new entrants and the number of students enrolled in first grade from the six to the eleven year old group.

For example, in Table A4 the 67,029 seven year old enrollment in first grade in 1979 was computed by multiplying the 1978 six year old population (111,716) (i.e., who were seven years old in 1979) by the enrollment rate for the seven year olds in 1979 as presented in Table A2 (60.0 percent). The 22,183

eight year old students enrolled in first grade in 1979 were calculated by multiplying the 1978 seven year old population (107,341) by the eight year old enrollment rate presented in Table A2 (20.7 percent). A similiar procedure was used to compute the nine, ten, and eleven year old enrollment. The total estimated number of students enrolled in first grade from the seven to eleven year old groups in 1978 by the model was 111,562. This number and the total number of new entrants were used to estimate the enrollment of the six year old group. The 36,709 six year old students enrolled in 1979 were computed as the difference between the 153,271 new entrants and the total enrollment of the seven to eleven year old students enrolled in 1979 (116,562). If the estimated number of six year old students are larger than the six year old population then the model assumes that all the six year old children come to first grade that year. This is what happened in 1983 when, according to the model, 132,308 six year old students were enrolled in first grade. Given that adding the total enrollment of the six to eleven year old group in 1983 gives a figure (132,796) that is lower than the number of new entrants for that year (173,616) the model uses the five year old population to make up for that difference. This is why in 1983 there are 40,820 five year olds enrolled in first grade.

Table A4: Simulation of Enrollment in First Grade For Intake Rates Larger Than 100% The Case of Honduras

	Enrollment in First Grade					
Age	1979	1980	1981	1982	1983	1984
5	0	0	0	0	40,820	43,297
6	36,709	51,733	79,832	127,357	132,308	136,680
7	67,029	68,887	67,581	43,626	488	0
8	22,183	13,406	9,841	0	0	0
9	12,391	11,092	0	0	0	0
10	7,291	6,196	0	0	0	0
11	7,668	3,646	0	0	0	0
New Entrants	153,271	154,959	157,254	170,983	173,616	179,977

For the first year that the simulation is run (in this example, 1979), the model assumes that there is no enrollment from the five year old group. Another assumption built into the model is that as time goes by each cohort runs out of people who have not attended first grade. This is the reason why, when these figures are used, an increasing number of the six and the five year old groups needs to enter the system each year to match the number of new entrants.

The result of running the simulation using data for Honduras is that the Honduran educational system would have needed to attract an increasing proportion of its six year old group for the number of new entrants to increase from 153,271 in 1979 to 179,977 in 1984. In fact the participation rate of the six year old group would need to be 100 percent (Table A3) by 1983, which is very difficult to believe. Also beginning in 1983 the system would have had to attract an increasing proportion of five year old children to the system. An intake rate of 179,997 students in 1984 would have needed to enroll 100 percent of the six year old population and 30 percent of the five year olds, both of which are very unlikely situations.

Bibliography

Ascher, William. "The Forecasting Potential of Complex Models." Policy Science, 13(3), 1981, pp. 247-267.

Becker, Gary. Human Capital. NY: Columbia University Press, 1964.

Berstecher, D. Costing Educational Wastage: A Pilot Simulation Study. Paris: UNESCO, 1970.

Bloomfield, Stefan and Updegrove, Daniel. "Modeling for Insight, not Numbers." EDUCOM Bulletin, 17(3), 1982, pp. 5-9.

Brimes, M. A. and Pauli, L. Wastage in Education: A World Problem. Paris: UNESCO. International Bureau of Education, 1971.

Cassidy, Thomas. "Microcomputer-based Models and Modeling Systems Software for Educational Planning." Qualifying paper, Harvard University, 1986.

Chapman, David W. and Boothroyd, Roger. "Threats to Data Quality in Developing Country Settings." Comparative Education Review, 32(4), pp. 416-429.

Crouch, L.; Moreland, S.; Cuadra, E.; and McGinn, N. A User's Guide to the Primary Education Tracking System. NC: RTI/Project BRIDGES, 1986.

Cuadra, Ernesto. "Development of a System for Tracking Primary Education in Central America: Final Report." Harvard University, Project BRIDGES, 1987. Photocopy.

Cuadra, Ernesto; Anderson, Mary; Moreland, Scott and Dall, Frank. "Female Access to Primary Education: Trends, Policies and Strategies." Harvard University, Project BRIDGES, 1988. Photocopy.

Davis, Russell G. Planning Human Resource Development. Educational Models and Schemata. Chicago: Rand McMilland, 1966.

Davis, Russell G. Planning Education for Development: Volume II. Models and Methods for Systematic Planning of Education. Cambridge, MA: USAID/HARVARD, 1980.

Dominguez-Urosa, Jose'. Efficiency and Other Indicators of Performance in Education Systems. Washington: World Bank, 1980.

Eshiwani, George. "Improving Access to Education: A Review of Published and Unpublished Research from Central and Southern Africa." Kenya: Bureau of Educational Research, Kenyata University, 1986.

Evanco, William. "Multinomial Logistic Enrollment Models: Application to Iraqi Primary Education." Educational Planning, 3(2), 1976, pp. 53-65.

Evans, David. "Micro-computer Modeling for Educational Planning and Policy Analysis. A Summary Outline of Key Issues." Paper presented at the 1986 meeting of the Comparative and International Education Society, Toronto, 1986.

Farrell, J. P. "National Planning Systems in Latin America: Their Environments and their Impact." Educational Planning, 1(1), 1974.

Farrell, J. P. and Schiefelbein, E. "Expanding the Scope of Educational Planning: The Experience of Chile". Interchange, 5(2), 1974, pp. 18-28.

Fiala, Robert and Lanford, Audri Gordon. "Educational Ideology and the World Educational Revolution." Comparative Education Review, 31(3), 1987, pp. 315-332.

Inkeles, Alex and Siroway, Larry. "Convergent and Divergent Trends in National Education Systems." Social Forces, 62, 1983, pp. 303-33.

International Bureau of Education (IBE). Wastage in Education: A World Problem. Paris: UNESCO/IBE, 1971.

International Bureau of Education (IBE). A Statistical Study of Wastage at School. Paris: UNESCO/IBE, 1972.

Jackson, G. B. "The Research Evidence on the Effects of Grade Retention." Review of Educational Research, 45, 1975, pp. 613-635.

Khoi, Le Thanh. "Planning Education: History." In Torsten Husen and T. Neville Postlethwaite (Eds.) The International Encyclopedia of Education. NY: Pergamon Press, 1985. Vol 7, pp. 3927-3929.

Kitch, D. E. "Does Retention Cause Dropouts?" California Journal of Elementary Education, 21, 1952, pp. 25-28.

Larkey, Patrick D. and Sproull, Lee S. "Models in Theory and Practice: Some Examples, Problems, and Prospects." Policy Science, 13(3), 1981, pp. 233-246.

Levy, M. B. "Determinants of Primary School Dropouts in Developing Countries." Comparative Education Review, 15(1), 1971, pp. 44-58.

Lockheed, M. and Hanushek, Eric. "Improving Educational Efficiency in Developing Countries: What Do We Know?" Compare, 18(1), 1988, pp. 21-37.

Murnane, Richard. "Improving Education Indicators and Economic Indicators: The Same Problems?" Educational Evaluation and Policy Analysis, 9, 1987, pp. 101-116.

Mannheim, Karl. Man and Society in an Age of Reconstruction. New York: Harcourt, 1949.

McGinn, Noel and Davis, Russell, G. Build a Mill, Build a City, Build a School. Cambridge, MA: MIT Press, 1969.

Meyer, John et al. "The World Educational Revolution, 1950-1970." Sociology of Education, 50, 1977, pp. 242-258.

Myers, Robert G. Connecting Worlds: A Survey of Development in Educational Research in Latin America. Ottowa: IDRC, 1981.

Oakes, J. Educational Indicators: A Guide for Policymakers. CA: The Rand Corporaton, 1986.

OECD. Indicators of Educational Performance. Paris: Organization for Economic Cooperation and Development (OECD), 1971.

OECD. A Framework for Educational Indicators. Paris: Organization for Economic Cooperation and Development (OECD), 1973.

Psacharopoulos, George and Woodhall, Maureen. Education for Development. An Analysis of Investment Choices. N.W.: Published for the World Bank by Oxford University Press, 1985.

Reynolds, H. T. Analysis of Nominal Data. Beverly Hills, CA.: Sage Pub. Series: Quantitative Applications in the Social Science, No. 7, second edition, 1984.

Sammak, A. "School Enrollment Coding System." Statistical Methods for Improving the Estimation of Repetition and Drop-Out: Two Methodological Studies. Paris: UNESCO, 1981, pp. 95-113.

Schiefelbein, Ernesto and Davis, Russell G. Development of Educational Planning Models and Application in the Chilean School Reform. Lexington, MA.: Lexington Books, 1974.

Schiefelbein, Ernesto. "Repeating: An Overlooked Problem in Latin American Education." Comparative Education Review, 19(3), 1975, pp. 468-487.

Schiefelbein, Ernesto. "M*todos Para Examinar Antecedentes Sobre Repeticion." La Educacion. A:o XXIII, No. 81, 1979, pp. 9-37.

Schiefelbein, Ernesto and Davis, Russell G. "Methods of Estimating Enrollment Flow Rates: An Instructional Unit." In Russell G. Davis, Planning Education for Development: Volume II. Models and Methods for Systematic Planning of Education. Cambridge, MA: USAID/HARVARD, 1980, pp. 121-149.

Schiefelbein, Ernesto and Grossi, M. C. "Statistical Report on Repetition in Latin America." Statistical Methods for Improving the Estimation of Repetition and Drop-Out: Two Methodological Studies. Paris: UNESCO, 1981, pp. 1-94.

Schiefelbein, Ernesto. "Planning Education in Latin America." In Torsten Hus*n and T. Neville Postleth-

waite (Eds.) The International Encyclopedia of Education. N.Y.: Pergamon Press, 1985, Vol. 7, pp. 3933-3936.

Schiefelbein, E.; Molfino, I.; Martinez, H.; and Grossi, M.C. Simulating Historical and Future Trends of an Educational System with the SMMG: The Case of Guatemala. Washington: World Bank, 1985.

Schultz, Theodore. "Investment in Human Capital." American Economic Review, 51, 1961.

Thonstad, Tore. Analyzing and Projecting School Enrollment in Developing Countries: A Manual of Methodology. Paris: UNESCO, 1980.

UNESCO. A System of Educational Indicators. Paris: UNESCO, ED.75/ WS 8, 1974.

Weiler, H. N. "Planning Education: An Overview." In Torsten Husen and T. Neville Postlethwaite (Eds.) The International Encyclopedia of Education. N.Y.: Pergamon Press, 1985. Vol 7 pp. 3923-3927.

Windham, Douglas M. Indicators of Educational Effectiveness and Efficiency. Florida: IEES, Florida State University, 1988.

White Plisko, Valena and Ginsburg, Alan. "Assessing National Data on Education." Education Evaluation and Policy Analysis, 8(1), 1986, pp. 1-16.

Methodology One: A New Approach to the Estimation of Enrollment Transition Rates: Estimating Rates with Minimal Data Availability

Introduction: The Problem

Knowing rates of repetition, dropping out, and entrance into first grade is of great importance to educational planners and policymakers in developing countries. Knowledge of these rates is essential to the accurate forecasting of enrollments and to the analysis of the internal efficiency of school systems. Standard methodological manuals explain how to calculate all these rates. (See, for example, Thonstad, 1980.) Unfortunately, the traditional methods of estimating these rates are fraught with serious problems. This paper suggests an alternative to the traditional method. The proposed alternative requires data available in all countries, and can be quickly computed using a microcomputer.

There are two fundamental problems with estimating repetition and dropout rates according to traditional methods. The first is the simplest: many demographically important countries (e.g., Egypt, Pakistan) simply do not report repeater or dropout data on any systematic basis. The second problem is that in many countries where the numbers of repeaters are reported, the numbers are consistently undercounted, and the amount of bias is usually unknown. Since the repetition rate is estimated as the number of students repeating a grade at time t divided by the enrollment in that grade at time t-1, the repetition rate will evidently be underestimated. Since the dropout rate is usually estimated as a residual, this rate will be overestimated. The rate of entrance into the first grade is also frequently estimated as a residual, and is thus also overestimated. The result will be that both policymakers and planners will be misled. Policymakers will tend to think that repetition is a less serious problem than it really is, and concentrate their policies on preventing dropping out. At the same time, they will be misled into thinking that the percentage of children entering school out of every population cohort is much higher than it really is, and they will tend to think the problem of access to schooling is close to being solved. Planners, on the other hand, if they use these rates to forecast enrollment, will tend to make forecasts that are too low, since they assume repetition rates to be lower, and dropping out rates to be higher than they really are.

A natural question then arises: could the rates of repetition and dropping out be estimated without reference to the absent or mistaken counts of repeaters? If we can assume that actual enrollment data are either better or at least consistently biased, are there some mathematical or statistical techniques that could be used which depend only on information about actual enrollments? Under these assumptions, a method of estimating repetition and dropping out rates that depended only on enrollment data might yield better estimates of these rates. Several logical alternatives immediately come to mind.

Extant Alternative Solutions

A first alternative is to think about solving systems of equations. Let's take first grade enrollment as an example, where the problem is that of estimating the first grade repetition rate and the rate of entrance into the school system. We know that the equation for first grade enrollment is

$$E_{1,t}=r_{1,t-1}E_{1,t-1} + e_tP_t$$

where $E_{1,t}$ is enrollment in first grade at time t, $r_{1,t-1}$ is the first grade repetition rate at time t-1, e_t is the rate of entrance of new students from the population of school entrance age into the first grade, and finally P_t is the population of school entrance age at time t. Now, under the assumption that the repetition and entrance rates do not change radically from year to year, and if we had three years of data on enrollment and two years of data on population of entrance age, we would have two equations in two unknowns,

$$E_{1,t}=r_1E_{1,t-1} + eP_t$$

$$E1_{,t-1}=r_1E_{1,t-2} + eP_{t-1}$$

where the unknowns are the repetition rate and the entrance rate. Note that we have removed the time

subscript from the rates to denote the fact that we are assuming the rates do not change from t-2 to t-1. The problem with this method is that the equations are so similar in slope that the solution tends to give results that are mathematically correct, but make no sense as enrollment transition rates. Take as an example the following data for first grade enrollment in two recent years in the Dominican Republic (all schools, both sexes):

Year	E_t	E_{t-1}	P_t
1	310,277	308,478	159,873
2	316,605	310,277	161,464

The first column represents the right-hand side of the equation above. The second column, representing last year's enrollment, would be multiplied by a repetition rate. The last column, representing this year's population, would be multiplied by the entrance rate. The solution to the two implicit equations in the table above, unfortunately, turns out to give a repetition rate of -2.54, and an entrance rate of 6.86. Obviously, this method cannot be offered as a general method. There are two potential problems with this method. First, it may seem that the key problem is the assumption that the rates do not change between the two years. As we will see later, however, this is actually a reasonable approximation. A second, more serious problem is that, given the assumption of exact equality of the rates in the two years, forcing exact equality in the relations determining enrollment, when the relations are so similar, yields estimated rates that can easily be in an "unreasonable" range. The advantage of the simultaneous equations approach, evidently, is that very few data are required; only three consecutive years of enrollment data and two consecutive years of population data.

A second alternative that seems intuitively appealing is to take longer time series and apply a "least-squares error" criterion to fitting the crucial rates. We would have a long time series of data on E_t, which would assume the role of "dependent" variable, and on E_{t-1} and P_t, which would assume the role of "independent" variables. We would then estimate our repetition and entrance rates as those which minimize the differences between the observed E_t and the E_t predicted with the estimated rates. This procedure would have the advantage of not requiring an exact fit with the observed data (i.e., it is more of a "statistical" rather than a "mathematical" approach), and so, hopefully, the estimated rates would make more sense. Unfortunately, it requires assuming that the rates do not change over a period much longer than two years, which is a much more restrictive assumption. Furthermore, the two "independent" variables (E_{t-1} and P_t) are highly correlated with each other, resulting in what statisticians term "multicollinearity", which means that the estimated rates will be quite unreliable from a statistical point of view. (As well as possibly meaningless from an educational point of view.) Finally, estimation of the transition rates using this method could be plagued by autocorrelation.

In response to these problems, educational planning researchers have been searching for alternative methods to estimating these rates. Schiefelbein (1980), for example, has developed various approaches, although (to varying degrees) those methods require more extensive data than the method we propose here.

A New, Simplified Approach

The approach proposed in this paper attempts to deal with the problems mentioned above. It does not require exact equality in the relations, as in the simultaneous equation option. It does not require a long time series of data, as in the least-squares option; nor does it require the shaky assumption that transition rates do not change over the longer period. It does not require data on enrollment by age and grade, which are not available in many countries. The method proposed has minimal data requirements, and the assumptions can be made as restrictive or as relaxed as the user feels comfortable with. Only three years of consecutive enrollment data, and two years of data on population of school entrance age are required.

The method has the advantage of mathematically maximizing the fit between the enrollment pattern produced by the estimated rates and the historical pattern in the three years of enrollment data. There are two disadvantages, however. The first is a methodological disadvantage; it may occasionally produce rates that seem unstable from grade to grade. (What this means will become apparent when we present an example.) Thus, the knowledge produced by this method may need to be tempered with knowledge derived from more traditional methods. In this sense, this method is offered not as an alternative to any others, but as a complement. The second disadvantage is that it is impossible to apply by hand; it requires access to a microcomputer and to linear

programming software. Fortunately, most planning and policy offices in ministries of education in developing countries now have access to microcomputer hardware, and getting linear programming software is not at all expensive.

What does the proposed method consist of? We saw above a description of the problem of estimating enrollment transition rates as a simultaneous equations problem. But we also saw that the straightforward use of simultaneous equation solutions led to results that are mathematically correct but make no sense from an educational point of view. What if we can take a modified simultaneous equation approach? Part of the problem, we saw, was that the "exactness" of the method was causing problems. What if one can relax the "exactness" aspect of simultaneous equations, and yet keep the "low data requirement" aspect? What if we could find a less "exact" method, e.g., one that could allow for inequalities? Furthermore, what if this method allowed one to incorporate restrictions that prevented results that might be mathematically correct but made no educational sense? One available method is to use the simplex algorithm of linear programming as a solution tool to maximize the fit between observed enrollment and the enrollment "predicted" by a set of estimated rates.

We pose the problem thus: we need to determine repetition rates for grades one through six, an entrance rate, and dropout rates for grades one through five. Thus, we have twelve unknowns. Are there twelve linear mathematical relations between these unknowns that we could use? If we assume that the rates do not change between a given two sets of two consecutive years, then we can have one fundamental enrollment equation for each of six grades, for each of two years: twelve relations. For any given year, and for any grade other than first grade (we have seen the first grade equation above), this equation is

$$E_{g,t} = r_{g,t-1}E_{g,t-1} + (1-r_{g-1,t-1}-d_{g-1,t-1})E_{g-1,t-1}$$

Let us suppose, however, that we do not insist on perfect equality as in the relation above, and at the same time assume that the rates do not change from t-2 to t-1. Let us call the right hand side of the equation above, with the estimated transition rates, the "predicted" enrollment, E*. Thus, we would have one equation

$$E^*_{g,t} = r^*_g E_{g,t-1} + (1-r^*_{g-1}-d^*_{g-1})E_{g-1,t-1}$$

for each grade in each of two years. Note that we have removed the time subscript from the rates, to indicate that they hold for both years. Now, the trick is to estimate transition rates so as to minimize the deviation the "predicted" values and the real values. Since there is no a priori reason to prefer a squared error criterion to an absolute error criterion, and since the latter is easier to handle within the context of linear programming, we minimize the sum of the absolute errors:

$$\text{Min } \Sigma_{g,t} \text{ abs } (E^*_{g,t} - E_{g,t}) \text{ over } r^*_g, d^*_g, \text{ and } e^*,$$

which is equivalent to

$$\text{Min } \Sigma_{g,t} E^*_{g,t} \text{ over } r^*_g, d^*_g, \text{ and } e^*,$$

subject to

$$E^*_{g,t} > E_{g,t}.$$

In addition, the following constraints need to be satisfied in order to define the relationship between the rates to be estimated and the predicted enrollment:

$$E^*_{1,t} = r^*_1 E_{1,t-1} + \qquad e^* Pt$$

$$E^*_{g,t} = r^*_g E_{g,t-1} + (1-r^*_{g-1}-d^*_{g-1})E_{g-1,t-1}$$

Furthermore, arbitrary but reasonable limits can be placed on various rates with nearly infinite flexibility and very explicitly, as long as these limits can be expressed linearly. For example, one might constrain the estimated entrance rate e^* so that $e^* < 1.0$, or some other reasonable upper bound. Note we have removed the time subscripts from the rates in the equations above, to signify that the same rates apply for both time periods under consideration.

All this can easily and quickly be solved with any linear programming package on a small microcomputer (e.g., What's Best), though it would be a true nightmare to try to do with a hand calculator. It could possibly be solved once with a hand calculator, but this would not be enough to allow the planner or policy analyst to "experiment" with the method. And it is this experimentation which is essential to the process of planning and policy analysis. With a microcomputer it takes literally minutes to set up and solve.

An Example

Suppose we have only the following data (national level data from the Dominican Republic, boys and girls combined):

	Population	Enrollment					
		Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6
Year 1	159,873	308,478	173,528	140,861	118,826	92,252	74,646
Year 2	161,464	310,277	177,423	145,347	123,951	97,173	79,687
Year 3	163,055	316,605	178,724	149,076	126,908	101,395	82,395

The problem then is, using only these data, to estimate the rates of transition between all the grades and between the population and the first grade.

The initial results with the constraints as listed above, are:

	Estimated transition rates		
	Entrance	Repetition	Dropout
Grade 1	1.0	0.49	0
Grade 2	—	0.12	0.13
Grade 3	—	0.12	0
Grade 4	—	0	0.21
Grade 5	—	0.03	0.10
Grade 6	—	0	—

Clearly there is a problem in that, as has been mentioned before, the rates are somewhat unstable from grade to grade. For example, it is unlikely that the repetition rates for grades four and six are really zero, or that the dropout rates for grades one and three are similarly zero. One of the advantages of the linear programming method as an approach is that the mechanical aspects of the methodology can be tempered with additional empirical information coming from many disparate sources. Suppose we have good reason to suspect that the rates for any given grade are not likely to differ by more than 50 percent from those of the next grade up or down, and that the average rates across all grades must be less than 30 percent but greater than 5 percent. Formally, this means adding the following constraining relations to the mathematical statement of the problem:

$r_g > 0.5\ r_{g-1}$

$r_g < 1.5\ r_{g-1}$

$d_g > 0.5\ d_{g-1}$

$d_g < 1.5\ d_{g-1}$

$\Sigma_g\ r_g < 0.3 * 6$

$\Sigma_g\ r_g > 0.05 * 6$

$\Sigma_g\ d_g < 0.3 * 6$

$\Sigma_g\ d_g > 0.05 * 6$

With these modifications, the method provides the following estimates of the transition rates, which seem much more reasonable, and yet are produced without using any official repetition data:

	Estimated (constrained) transition rates		
	Entrance	Repetition	Dropout
Grade 1	1.0	0.49	0.07
Grade 2	—	0.25	0.05
Grade 3	—	0.16	0.06
Grade 4	—	0.12	0.11
Grade 5	—	0.06	0.10
Grade 6	—	0.03	—

Adding any assumptions and prior information can thus temper the mechanical results produced by the method. Fortunately, the assumptions can be expressed in a very clear and explicit manner, and the effects of the assumptions on the estimated rates are easy to test. Note, for example, that the estimated repetition rate for the first grade has barely changed, and that repetition rates as estimated here are consistently higher than dropout rates, even with the

new assumptions. The method allows the analyst to rapidly test the effects of many sets of reasonable restrictions. If certain results are robust to the various sets of reasonable assumptions (for example, if the first grade repetition rate remains between 0.4 and 0.6 under all reasonable scenarios), then it is safe to conclude that those results are likely to be a true reflection of the underlying rate structure.

Conclusion

The method presented in this paper has the capability of offering the analyst an alternative to traditional methods of estimating repetition and dropout rates. This alternative method relies only on widely available data, and can explicitly incorporate any prior knowledge about the rates. Evidently, it is not an exact method. But, in combination with other available methods, it can be used to try to get a closer fix on what the true repetition and dropout rates are in many developing countries.

References

Schiefelbein, Ernesto, and M.C. Grossi. 1980. Statistical Report on Repetition in Latin America. UNESCO, Division of Statistics on Education, Office of Statistics.

Thonstad, Tore. 1980. Analysing and Projecting School Enrollment in Developing Countries: A Manual of Methodology. UNESCO, Statistical Reports and Studies Series No. 24.

What's Best, General Optimization, Inc., Chicago.

Luis Crouch
Research Triangle Institute
BRIDGES Project

Work on this paper was funded by the Basic Research and Implementation for DevelopinG Education Systems (BRIDGES) project, which is sponsored by the U.S. Agency for International Development, Office of Education, under Cooperative Agreement No. DPE-5824-A-5076.

Methodology Two: Monitoring Internal Efficiency of a School System in the Absence of Reliable Repeater and Dropout Data: A Method Based On Simple Proxy Indicators

The Problem

School systems in many countries do not report repeater and dropout data. In those countries where these data are reported, the information is often of dubious reliability. Yet both repeater and dropout data are essential to the monitoring of changes in internal efficiency of the system, so that authorities may know "what works," or indeed, how well the system is working. Perhaps the most commonly used monitoring indicator is the primary school completion rate. But, as normally calculated, this rate is the result of analyzing the school history of a "reconstructed" or hypothetical cohort of 1000 entrants into the system. This school history is "reconstructed" using repeater and dropout rates. Thus, when these rates do not exist, or are of doubtful quality, the primary school completion rate cannot be known, or cannot be relied upon. Naturally, any method that can give one an approximate idea as to whether the completion rate is improving or not would be a welcome addition to the "toolkit" of educational planners and education sector analysts. This note proposes a method that is very easy to use and requires no data on repeaters and dropouts.

A Proposed Proxy Method[1]

In what follows we will assume a five-grade school system, with students entering at age five and reaching fifth grade at age nine, but all our conclusions apply to any system.

The proxy indicator of the completion ratio proposed here is a proportion of two ratios: the ratio of enrollment in the last grade to the population cohort of corresponding age, (call it the last grade enrollment ratio), divided by the growth-adjusted overall gross enrollment ratio.[2] In a system where this proportion is increasing, we may conclude that the completion rate is increasing.

Why a growth-adjusted gross enrollment ratio? In a system that is both stable (not growing) and perfectly efficient, the fifth grade enrollment ratio would be exactly the same as the overall enrollment ratio, since 100% of the children would flow from one grade to the next, and there would be no new "catch-up" entrants into the first grade of newly opened schools. Thus, the number of children in the first-grade grade would be the same as the number of children in the first grade, and the last-grade enrollment ratio and the overall gross enrollment ratio would be exactly the same. However, a system could be very efficient, in that there are few dropouts and repeaters, and yet have a last-grade enrollment ratio which is lower than the overall enrollment ratio, due to the expansion of the system (which naturally produces a bulge in the enrollment figures of the earlier grades). In practice, for any real system typical of most LDCs, the difference between the overall enrollment ratio and the fifth grade enrollment ratio can be attributed either to inefficiency or to growth, and the growth could be due to either expansion of the school system into rural areas, or simply fast population growth. If one could find a way to remove the growth effect from the overall gross enrollment ratio, then the remaining effect is inefficiency.

How To Adjust The Gross Enrollment Ratio For System Growth

Let's take as an example a five-grade primary system. If the fifth grade enrollment ratio at time t can be compared with the first grade enrollment ratio at time t-4, then much of the effect of growth can be removed. Better yet, if we take the weighted average of the first grade enrollment ratio at time t-4, second grade enrollment ratio at time t-3, and so on, and compare this to the fifth grade enrollment ratio at time t, we will have a pretty good idea of the effect of inefficiency change by itself.

Here are some examples of the calculations. Let's use boys' enrollment as an example this time. First, note the population figures:

Population of Boys (in 1000s):

Age	1983	1984	1985	1986	1987
5	72	74	76	77	89
6	70	72	73	75	77
7	68	69	71	73	75
8	66	67	69	71	73
9	64	65	67	69	71

Enrollment of Boys (in 1000s):

Grade	1983	1984	1985	1986	1987
1	87	97	121	146	159
2	34	40	51	48	67
3	25	28	35	35	38
4	19	21	27	27	30
5	16	16	18	19	22

The growth-adjusted gross enrollment ratio for 1987 would therefore be (22+27+35+40+87)/(71+71+71+72+72), which yields 59.1%. Now, the fifth grade enrollment ratio for 1987 would be simply 22/71, which yields 31%. The ratio of 31 to 59.1 is 52.5%, and this is the indicator we are interested in calculating. If this number goes up, we can conclude efficiency is increasing. Note that the non-adjusted, or standard gross enrollment ratio for 1987 would be (22+30+38+67+159)/(71+73+75+77+89), which is 82.1.

The interpretation of these numbers is pretty straightforward. Note again that the last-grade enrollment ratio is about 31, whereas the gross enrollment ratio is 82. It would therefore seem that about 50% of children are being lost between first and fifth grades. But the relevant gross enrollment ratio is the growth adjusted gross enrollment ratio, which is only 59. Under this scenario, only about 30% of the children appear to be lost. This is a more realistic picture. Note that the factor that contributes to this difference is the tremendous expansion in first grade enrollment, due to a fast growing population and a school system that is expanding into rural zones. It is evident that when first grade enrollment is expanding so quickly it is inappropriate to compare the fifth grade enrollment ratio in any given year to the overall enrollment ratio for that same year. Instead, our proposed method compares fifth grade enrollment to the growth-adjusted total enrollment ratio, giving a more accurate picture.

Appendix. Arithmetical formulas.

The proposed indicator is

$GER_{5,t} / GER^*_t$

where

GER_5 is the fifth-grade enrollment ratio at time t, and

GER^*_t is the growth-adjusted gross enrollment ratio at time t.

The equation for GER_5 is

$$GER_5 = E_5 / P_9$$

where

E_5 = fifth grade enrollment, and

P_9 = population of nine year olds.

The equation for GER^*_t is

$$GER^*_t = (E_{5,t} + E_{4,t-1} + E_{3,t-2} + E_{2,t-3} + E_{1,t-4}) / (P_{9,t} + P_{8,t-1} + P_{7,t-2} + P_{6,t-3} + P_{5,t-4})$$

Endnotes

1. A simpler alternative to the proposed method is the ratio of current enrollment in the last grade to enrollment in the first grade six years before. The problem with this method is the inherent bias due to population growth, which would lead to overestimates of system inefficiency.

2. See the arithmetical statement in the appendix.

Luis Crouch
Research Triangle Institute
BRIDGES Project

Work on this paper was funded by the Basic Research and Implementation for DevelopinG Education Systems (BRIDGES) project, which is sponsored by the U.S. Agency for International Development, Office of Education, under Cooperative Agreement No. DPE-5824-A-5076.

BRIDGES Publications Staff:

Billie Jo Joy, Publications Coordinator
Pamela Summa, Editorial Assistant
Brenda Salyer, Graphic Designer

For more information about this or other BRIDGES documents, please contact:

BRIDGES Publications
442 Gutman Library
Harvard University
Cambridge, MA 02138
USA
(617) 495-9720

or

USAID
Office of Education
Bureau for Science and Technology
Department of State
Washington, D.C. 20523 USA
(703) 235-8980

A program of:
USAID, Bureau for Science & Technology

Directed by:
The Harvard University Institute for International Development
and the Harvard Graduate School of Education

In Cooperation with:
Academy for Educational Planning and Management, Pakistan
Centre de Perfectionnement et de Formation en Cours d'Emploi, Burundi
Institute for International Research
Michigan State University
The Open University, Indonesia
Research Triangle Institute
Texas Southern University
National Education Commission of Thailand
National Institute of Education, Sri Lanka
USAID Missions and Ministries of Education in Burundi, Egypt, Indonesia, North Yemen, Pakistan, Sri Lanka, and Thailand